João Manuel da Costa e Cruz

Licenciado em Ciências da Engenharia Electrotécnica e de Computadores

Context-based Identification of Energy Consumption in Industrial Plants

Dissertação para obtenção do Grau de Mestre em
Engenharia Electrotécnica e de Computadores

Orientador : Prof. Doutor Rui Neves-Silva, Professor Auxiliar,
Universidade Nova de Lisboa

Júri:

Presidente: Prof. Doutor Paulo José Carrilho de Sousa Gil

Arguente: Prof. Doutora Maria Júlia Fonseca de Seixas

Vogal: Prof. Doutor Rui Neves-Silva

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Dezembro, 2014

João Manuel da Costa e Cruz

Graduated Studies in Electrical and Computer Engineering

Context-based Identification of Energy Consumption in Industrial Plants

Dissertation to obtain the Master Degree in
Electrical and Computer Engineering

Supervisor: Prof. Rui Neves-Silva, PhD, Assistant Professor,
Universidade Nova de Lisboa

Evaluation Board:

President: Prof. Paulo José Carrilho de Sousa Gil, PhD

Opponent: Prof. Maria Júlia Fonseca de Seixas, PhD

Member: Prof. Rui Neves-Silva, PhD

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

December, 2014

Context-based Identification of Energy Consumption in Industrial Plants

To my Professors, Family and Friends.

Acknowledgements

I take this space to leave my sincere thanks to all who contributed to the accomplishment of this work.

To Professor Rui Neves-Silva, who I greatly admire, for his availability, guidance and support through all the stages of this work.

To my Professors Luís Brito Palma and Fernando Coito, from who I learned a lot, for all the knowledge shared.

To my colleagues, André Águas, Daniel Marques, Ricardo Mota and Sérgio Coelho for their continuous motivation, but most of all, for their friendship and all the moments we shared together

To Eng. Tiago Gonçalves, CEO of InnoWave Technologies, for his understanding and motivation given during the last stage of my dissertation

To my family from who I'm eternally grateful, who always give me support and happiness. I also take this opportunity to express my sincere apologies, for all the time I was absent.

To Ana Pimentel for all her patience, motivation and for sharing her life with me.

I also want to express my gratitude to the project LifeSaver - Context sensitive monitoring of energy consumption to support energy savings and emissions trading in industry (G.A. FP7-ICT-287652), which motivated this line of work.

"If I have seen further than others, it is by standing upon the shoulders of giants."

– Sir Isaac Newton

Abstract

Nowadays, reducing energy consumption is one of the highest priorities and biggest challenges faced worldwide and in particular in the industrial sector. Given the increasing trend of consumption and the current economical crisis, identifying cost reductions on the most energy-intensive sectors has become one of the main concerns among companies and researchers.

Particularly in industrial environments, energy consumption is affected by several factors, namely production factors(e.g. equipments), human (e.g. operators experience), environmental (e.g. temperature), among others, which influence the way of how energy is used across the plant. Therefore, several approaches for identifying consumption causes have been suggested and discussed. However, the existing methods only provide guidelines for energy consumption and have shown difficulties in explaining certain energy consumption patterns due to the lack of structure to incorporate context influence, hence are not able to track down the causes of consumption to a process level, where optimization measures can actually take place.

This dissertation proposes a new approach to tackle this issue, by on-line estimation of context-based energy consumption models, which are able to map operating context to consumption patterns. Context identification is performed by regression tree algorithms. Energy consumption estimation is achieved by means of a multi-model architecture using multiple RLS algorithms, locally estimated for each operating context.

Lastly, the proposed approach is applied to a real cement plant grinding circuit. Experimental results prove the viability of the overall system, regarding both automatic context identification and energy consumption estimation.

Keywords: Energy consumption; Context awareness; Regression tree; Multi-models; Recursive least-squares.

Resumo

Atualmente, um dos grandes desafios a nível mundial e em particular no setor industrial, é reduzir o consumo energético. Dada a atual crise económica e o contínuo aumento do consumo, a identificação de possíveis situações de redução de custo tem-se tornado uma das principais preocupações entre empresas e investigadores.

Em meio industrial, o consumo energético é afetado por vários fatores, nomeadamente fatores de produção (e.g. equipamentos), humanos (e.g. experiência do operador), ambientais (e.g. temperatura), entre outros, que influenciam a forma como a energia é usada pela instalação. Com o objetivo de identificar como estes fatores influenciam o consumo, várias metodologias têm sido sugeridas. No entanto, estas apenas fornecem diretrizes, e têm demonstrado dificuldade em explicar determinados padrões de consumo devido a não incorporarem uma estrutura que inclua a influência do contexto. Por este motivo, estas técnicas não são capazes rastrear as causas do consumo energético até ao nível do processo, onde efetivamente medidas de otimização podem ser tomadas.

Esta dissertação propõe uma nova abordagem para superar este problema, utilizando modelos de consumo energético baseados em contexto, que relacionam o contexto de operação com os padrões de consumo energético. A identificação de contexto é efetuada através de algoritmos de árvores de regressão e o consumo é estimado através de uma arquitetura baseada em múltiplos modelos locais, por meio de múltiplos algoritmos RLS, um por cada contexto de operação.

Por último, a abordagem proposta é aplicada a um circuito de moagem de cimento. Os resultados experimentais comprovam que a metodologia desenvolvida é viável, tanto na identificação automática de contextos como na estimação do consumo energético.

Palavras-chave: Consumo Energético; Identificação de Contexto; Árvores de regressão; Múltiplos modelos; Mínimos quadrados recursivos.

Contents

List of Figures

List of Tables

List of Symbols

y	Energy Consumption.
u	Production Variable.
f	Mapping Function.
e	Disturbances.
w	Context Variable.
R	Context Region.
c	Centroid.
ξ	Partition Error.
τ	Minimum Period of Steadiness.
κ	Quadratic Error Tolerance.
ε	Quadratic Error Threshold.
$\hat{f}$	Mapping Function Approximation.
Θ	Estimated Vector of Parameters.
t_k	Discrete Time.
T_s	Sampling Period.
φ	Regression Vector.
$\hat{y}$	Energy Consumption Estimation.
ϵ	Estimation Error.
$\hat{f_M}$	Local Mapping Function Approximation.

Θ_M	Local Estimated Vector of Parameters.
$\hat{y}_M$	Local Energy Consumption Estimation.
J	Cost Function.
P_M	Local Model Covariance Matrix.
λ	Forgetting Factor.
N_λ	Forgetting Window.
$\overline{\lvert\Delta\epsilon\rvert}$	Average Absolute Estimation Error.
N	Vector Length.
γ	Bank Sensitivity.
F_c	Flow of Clinker.
F_g	Flow of Gypsum.
F_l	Flow of Lime.
F_t	Flow of Trass.
F_{ce}	Flow of Cement.
P_{drive}	Drive Power.
y_{sc}	Specific Consumption.
Q	Ratio of Clinker.
F_{total}	Total Flow.
$\hat{y}_{sc}$	Specific Consumption Prediction.

Acronyms

FFANN	Feed-forward Artificial Neural Network.
LSRT	Least Squares Regression Tree.
CBEM	Context-based Energy Models.
RLS	Recursive Least Squares.
EIA	U.S. Energy Information Administration.
SEC	Specific Consumption.
EUP	Energy Use Parameters.
AI	Artificial Intelligence.
SVM	Support Vector Machines.
L SVM	Linear Support Vector Machines.
NL SVM	Non-Linear Support Vector Machines.
LL SVM	Locally Linear Support Vector Machines.
LNL SVM	Locally Non-Linear Support Vector Machines.
ANN	Artificial Neural Network.
GA	Genetic Algorithm.
FS	Fuzzy Systems.
RT	Regression Tree.
SSR	Sum of Squared Residuals.

1

Introduction

1.1 Context and Motivation

In recent years, competition in the industrial world has become more intense. As a result, the necessity for finding new ways or optimizing energy consumption has increased.

To facilitate better planning, companies often maintain databases that store energy consumption measurements along with usage patterns and production data of their major appliances. These databases are then used to build energy consumption models which help the estimation of energy demand and the identification of possible costs reductions.

However, this conventional models, have shown difficulties in explaining certain energy consumption patterns, mostly due to the fact that energy consumption in industrial plants depend on multiple factors, namely, production factors (e.g. production rates, type of product), human factors (e.g. operator experience), ambient factors (e.g. temperature), among others, and the sum of all of these factors affect the way of how energy is used.

Another driving factor is the recent rapid development of new technologies, which allow the acquisition of information from new sources. Particularly, dispersed contextual information can be acquired by means of dispersed intelligent sensor technology - Ambient Intelligence (AmI). These new technologies provide the ability of measuring dispersed contextual information besides the conventional production data, allowing companies to build energy consumption models based simultaneously on contextual and production information.

Not only acquiring information has become easier, handling large amounts of data has also been improved by the development of data mining and alternative approaches of estimation methods.

The rise of these new technologies along with the combination of what they provide, will bring companies new opportunities for optimization consumption, driving this work to the study of a method to develop context-based energy consumption models.

1.2 Objectives and Original Contributions

The overall objective of this dissertation is to develop a method to support companies in optimizing their operations and energy savings, by being able to identify how energy is used and employed across the several stages that comprise the company's production chain.

When process data retrieved from production unit operations is put together with energy consumption information and contextualized into operating contexts, a better understanding of how energy flows through the several production lines is achieved. This work, proposes a new generic and adaptive method for learning energy consumption patterns observed through the production stages, based on the contextual conditions in which production units are operating.

Identifying context-based energy models brings two significant contributions.

The first and most evident, is that more accurate estimations of energy consumption can be performed.

The second is the better understanding of how context differently impacts energy consumption, helping companies on finding alternative production strategies.

Both, significantly contribute to better ways of using energy, reducing not only costs but also improving companies overall efficiency.

1.3 Dissertation Organization

This dissertation is organized on 7 chapters (including this one) and 3 appendixes.
The rest is organized as follows:

- **Chapter 2 - State of Art** — It is approached the thematic of energy consumption in industrial plants, optimization methods and some related works. State of the art consumption models and major applications are also analysed.

- **Chapter 3 - Context Identification** — It is presented the concept of context variable and how it impacts energy consumption. A regression tree algorithm is proposed for context identification followed by computational simulations for theoretical demonstration and also for real case scenario framework.

- **Chapter 4 - Context-based Energy Models** — Is presented the concept of context-based identification of energy consumption, starting from the traditional estimation method and ramping to the context-based multi-model approach. A method to relate context to models is also presented alongside with computational simulations.

- **Chapter 5 - Implementation** — It is depicted the context identification method incorporated in the multi-model structure and how they play together to build the overall context-based energy models. The description of a simulator developed for demonstration purposes and real case scenarios is also described.

- **Chapter 6 - Experimental Results** — It is analysed the results of the application in a real case of a cement plant manufacturing process, focused on the cement mill energy consumption, comprising plant acquired data analysis and results of the developed system. As the experiment was performed for a 5 month mill operation, only 3 months (September 2012, October 2012 and January 2013) with the most relevant data are described in detail. The other 2 months (November and December 2012) are briefly described in Appendix A.1 and A.2.

- **Chapter 7 - Conclusion** — The main conclusions taken from the developed system analysis are highlighted and further work paths are suggested.

- **Appendix A.1 - Cement Mill Data of November and December 2012** — It is depicted the acquired data and a brief analysis for the months of November and December 2012.

- **Appendix A.2 - CBEM Experimental Results of November and December 2012** — The performance of the developed system is briefly analysed for the months of November and December 2012.

- **Appendix A.3 - Publish Paper** — Attached the published article resultant from this dissertation.

2

State of Art

In this chapter is performed a brief overview of energy consumption in the industrial sector, starting from current global energy use, narrowing down to the more intensive-energy industrial sectors and main production units.

A top-down approach to keep track of how energy is used across an entire factory is present, along with current methods for optimizing consumption.

Current energy modelling technologies are also described, highlighting its applications, advantages and disadvantages.

Lastly a brief description of this work approach is given, and how it contributes to the overall energy consumption optimization process.

2.1 Energy Consumption in Industrial Plants

2.1.1 Global Energy Use Perspective

According to the U.S. Energy Information Administration (EIA), the industrial sector consumes about 50% of the world's total delivered energy [1]. (Figure 2.1).

The industrial sector is definitely the most energy-intensive sector when compared to any of the others, and is expected to grow from 200 quadrillion BTU[1] in 2010 to 307 quadrillion BTU in 2040, increasing by an average of 1.4% per year [1].

The worldwide industry sector comprises several sub-sectors, namely the manufacturing industry (e.g. chemical and petrochemical, iron and steel, non-metallic minerals, textile and others), agriculture, mining, etc [2, 3]. Within these industries, the largest part of energy consumption is accounted by the manufacturing sector, in which the chemical

[1] BTU is the British Thermal Unit and is a traditional unit of energy equal to about 1055 joules.

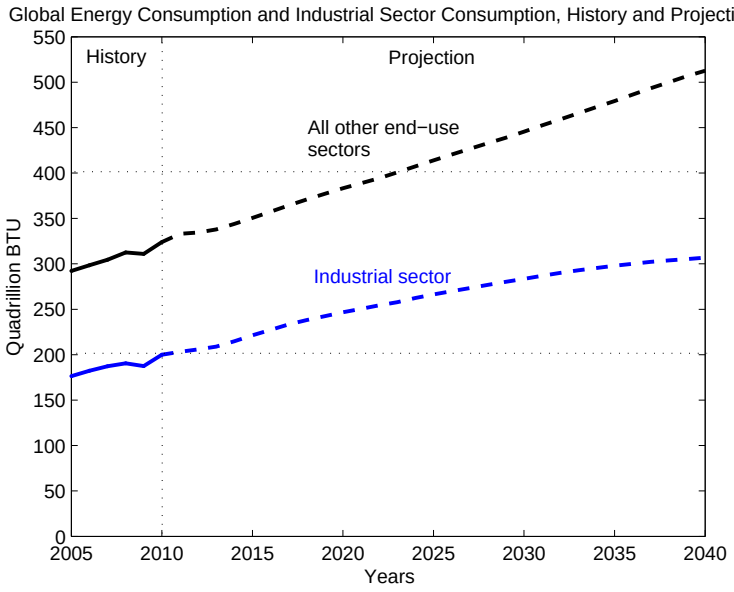

Figure 2.1: Global industrial sector and all other end-use energy consumption sectors from 2005 to 2040. Adapted from [1].

industry comprises a total of nearly 19%. The second largest energy consumer is the iron and steel industry, which accounts for 15%, followed by the non-metallic minerals sector with about 7%. The food and tobacco along with a large number of other categories of industrial energy users, account for the remaining 43.1% as depicted in figure 2.2 [1].

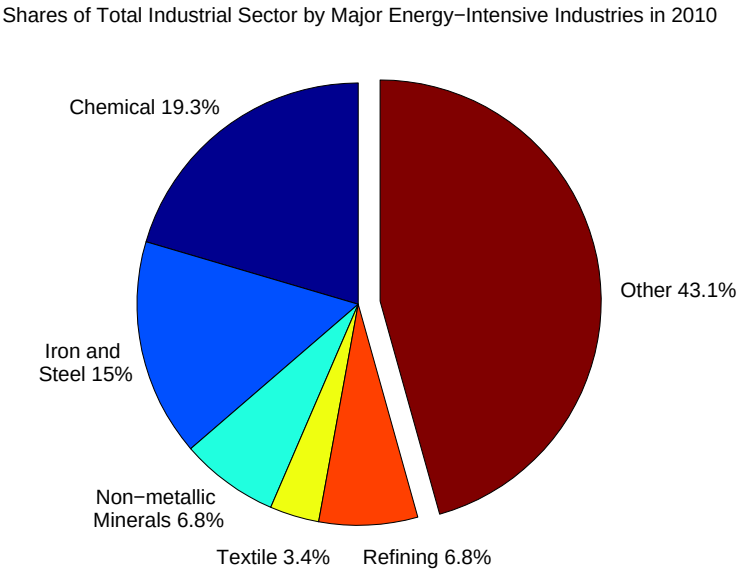

Figure 2.2: Shares of total industrial sector, by the major energy-intensive sectors in 2010. Adapted from [1].

Focusing on the manufacturing sector, a great part of the final energy demand is dependent on the processes that are allocated to perform specific materials production. Each industry comprises several unit operations ranging between simple milling processes to complex multi-phase high-temperature pyroprocessing [2] [4].

Following this research line and for brevity but without loss of generality, consider an overview of a cement plant energy consumption, highlighting the main energy-consumer processes (figure 2.3), which will be the focus of the practical study performed further on.

Grinding and pyroprocessing are the major production unit in a cement plant and are the most energy-intensive and inefficient operations accounting for a significant amount of the total energy consumed. These processes are multi-variable, time-variable and non-linear where complicated physical and chemical reactions take place [5].

Considering such a high energy consumption efficiency drop, particularly in the cement mill grinding circuits, small increases lead to significant reductions of energy consumption and obviously production costs. [6] Therefore, optimization of cement grinding circuits is much interesting and will be the object of study later on this work.

Because of the current global energy crisis and the impact of the industrial sector in the global energy consumption, whose trend is to increasingly use more energy, optimizing consumption has become one of the top research priorities [2].

Hence, this research line drives the current study to the next section, where state of the art methods for optimizing energy consumption are depicted.

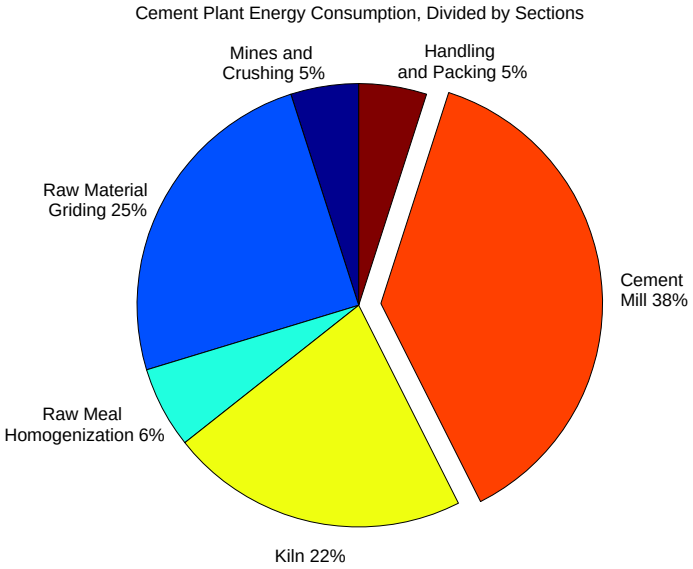

Figure 2.3: Cement plant energy consumption, divided by the energy-consuming processes. Adapted from [4].

[2] Pyroprocessing is a process in which materials are exposed to high temperatures, resulting in chemical or physical changes. Furnaces and kilns are the most used equipments for pyroprocessing.

2.1.2 Energy Consumption Optimization

In literature, there are several different approaches for optimizing energy consumption, which differ according to industrial sector, processes, equipments, etc. However, they all have in common a basic method to derive decisions on energy optimization measures (Figure 2.4), which comprise the following steps [7]:

1. Gather information about the process;
2. Model and analyse the process energy consumption based on the information gathered, finding relations between consumption patterns, production data and other significant variables that may impact consumption;
3. Determine which optimization measures should be taken according to the analysis performed;
4. Implement and validate optimization measurements.

The first step is gathering of information about the process main variables, process behaviour and influences from an energetic context, along with energy consumption patterns.

Based on this information, analysis are performed to define the current process operating regime from an energy consumption point view. Hence, detailed measurements, ideally acquired for each individual process, should provide enough information so that further energy consumption analysis are able to identify the causes of consumption.

Next is the consumption analysis step. In literature it can be found several ways for analysing consumption, which can be divided into 4 different types of approaches as described in [7, 8]:

- **Indicator-based** — Energy consumption analysis is based on information regarding indicators that define the consumption efficiency, namely economic, environmental

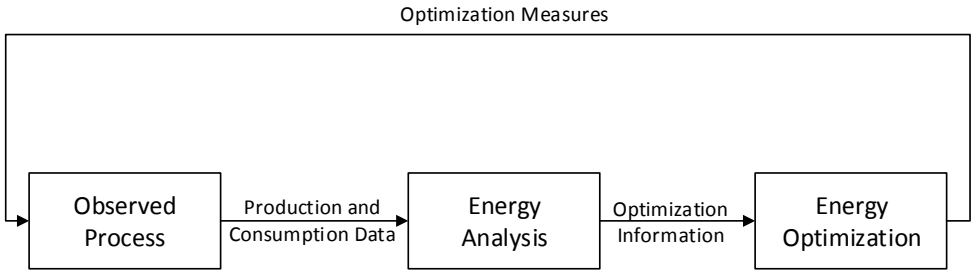

Figure 2.4: Basic energy consumption optimization process. Adapted from [7].

.

and physical [9]. The most used indicator in the industrial sector is the ration between a unit of final product and the amount of energy used to produce it, defined as specific consumption SEC expressed in kW/t [10];

- **Knowledge-based** — Energy consumption analysis is made based on knowledge either by human experts or by specialized automatic systems in a form that ranges between pre-defined lists of steps and by industrial best practices (in case of humans) or diagnostic and fault detection systems [11, 12];
- **Resource-based** — This approach is focused on the best way to deliver energy from a source to a consumer system, taking into consideration optimal grid stability, distributed generation, distributed storage ad demand side load management [13, 14];
- **Simulation-based** — Energy consumption is analysed by means of models which simulate plant dynamic. This models are then used to perform simulations that allow the identification of causes impacting energy consumption or even to examine the structure of the plant and how it behaves when structural changes are made [15].

Independently of the approach taken, the analysis should result in optimization measures that are able to eliminate the cause of consumption. However, sometimes this process is constraint by implementation costs, technical aspects and external side conditions.

Lastly, the solution found is implemented for the observed processes. At this stage, existing consumption models may be used to analyse and validate if the optimized solution is feasible and also help on defining the implementation steps.

The overall optimization process is dependent of the information gathered about the process operation and how it is correlated to certain energy consumption patterns so that possible causes that may impact energy consumption are identified. Hence, a structure to keep track of how energy is used is required, which leads us to the next section regarding how energy consumption is monitored in industrial environments.

2.1.3 Keeping Track of Energy Consumption

The well understanding of how and where energy is used across the plant is required, in order to identify possible opportunities to improve/reduce energy consumption, allowing a better planning of plant's operations and even the formulation of alternative production strategies.

This process normally follows a top-down approach as depicted in figure 2.5, comprising 3 major steps [16]:

1. First is the division of the overall plant into departments and delineate each department range of activity and responsibilities. Each department should then identify within its range of activity, which are the processes that consume more energy and should be followed closely. Also, the equipment and people organization, along

with facility information, such as air conditioning, lighting, and heating should be put together in order to produce an energy efficient workflow and resource assignment [2];

2. The second step is to monitor each individual process at each stage of the production chain to keep track of the energy use and main variables of process activity. Within a production chain, processes such as: grinding, milling, pyroprocessing, among others may comprise unique mechanical, chemical procedures. Hence the selection of variables should be carefully selected for each process [2];

3. Lastly, production rate is required to keep track of the specific consumption at each phase of the production. SEC sets a relation between the amount of product produced and the energy consumed, hence allowing comparisons between processes in the same production chain or even across each departments.

It is of extremely importance to keep track of energy consumption at process level, given that an assessment taking only into consideration a department consumption is not accurate enough to derive possible causes of consumption, due to the complexity and overall interactions of some processes [3].

Hence, measurements and on-line monitoring are fundamental for the acknowledgement of how processes behave and identification of energy consumption causes, particularly because processes are time varying and highly non-linear. Several technologies

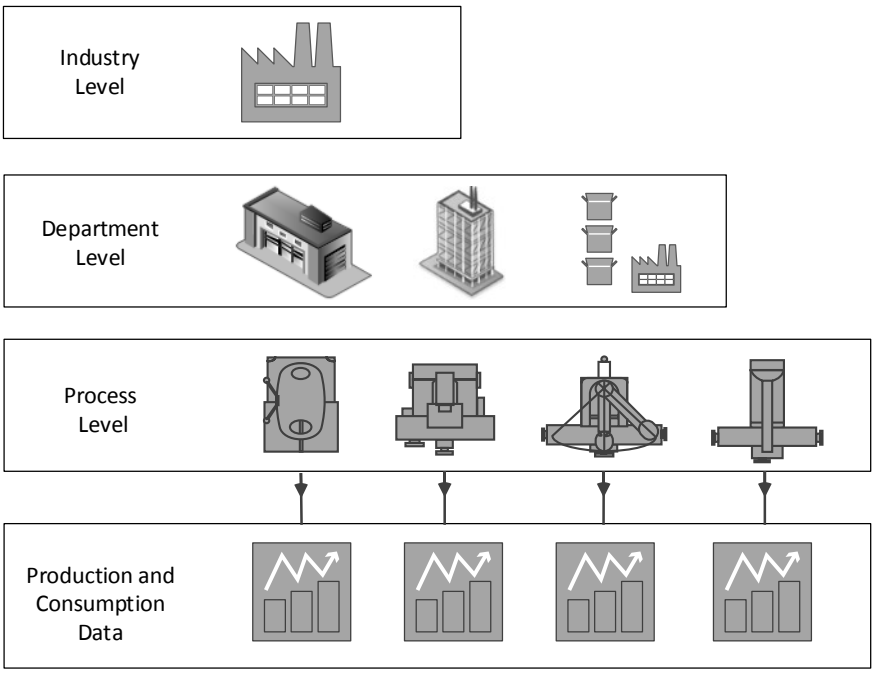

Figure 2.5: Energy track in industrial environments.

have been used to acquire real-time energy and process variables information, using distributed sensors and meters over wireless networks [10, 17].

The decomposition of the overall consumption into hierarchical level follows a concept of Energy Use Parameters (EUP). EUP derives from the acquire data to identify the use of energy as part of the facility, departments and processes, relating energy consumption data with production factors, such as: equipment type and use, production contextual conditions, production rates, product types under production, human interactions, process tasks, among others. There is no global rule that can be applied to set the number of levels in EUP, hence the several variations of this method [10, 18].

By putting together energy consumption data, process data and other variables that may impact consumption, process operation or even the overall consumption of each department, is possible to derive the causes of excessive consumption, which lead us to next section of this study, concerning main causes and influences on energy consumption in industrial environments.

2.1.4 Causes and Influences on Energy Consumption

In industrial environments, when the plant is performing a specific task, it is often affected by certain operating conditions that are dependent on multiple factors, namely, production factors (e.g. type of machines and equipment, type of product under produced, raw material quality), human factors (e.g. operators experience), ambient factors (e.g. temperature, humidity), among others, which impact the way of how energy is used, leaving certain energy consumption footprints [19].

One type of influence is given by the operators that interact with the plant. These interactions often affect the current task being performed, e.g. due to the operator experience on handling the plant equipment [7].

Another type of influence is the environment. Certain processes (e.g. pyroprocesses, heating, cooling, among others) that comprise thermodynamic and thermochemical reactions have strict temperature operating points. Thus, environmental conditions (e.g. ambient temperature and humidity) have a strong impact on the energy needed to keep the plant operation under a desirable operating state [20, 21].

In return, certain industrial tasks (e.g. welding) may impact the surrounding environment. In this case, not only temperature may be affected but also air quality and levels of brightness. Hence, for maintaining the work conditions under a normal state, more energy is spent in HVAC, ventilation and lighting systems [19].

Also, materials quality used for production may impact energy consumption. Often, many manufactures tend to replace materials by cheaper ones with similar properties, in order to reduce production costs from an economical point of view. However, this approach sometimes brings consequences at process level. A clear example is the replacement of clinker[3] by slag[4] in the cement industry. From a short-term economic point

[3]Clinker is the essential ingredient on cement production.

[4]Slag is a broad term comprising all non metallic products resulting from the separation of a metal from

of view the usage of slag instead of clinker makes sense, as it is much cheaper to acquire. However from a long-term point of view, it is clear inefficient as slag is harder to grind, increasing the SEC of the grinding circuit, thus requiring more energy to produce the same amount of cement [22].

Although the diversity of causes and influences on energy consumption, they must be diagnosed so that optimization processes have enough information to determine optimization measures.

Hence, models structure must account for this external influences, in order to capture and map the impact of all these factors into observed energy consumption patterns. This research line, drives the current study to the next section, where state of the art energy consumption models are described.

2.2 Energy Consumption Models

Energy consumption models have several applications on industrial environments. The first and most obvious is the ability to forecast future energy demand. However, energy models are also a way to understand how energy is used through production chains and main departments or to simulate how a plant would react to certain changes that wouldn't be possible to assess in real experiments.

Several research lines about modelling energy consumption have been discussed over the years, leading to a wide range of methods and different types of models.

Some are focused on building energy models based on the physical behaviour and process dynamics, extracting mathematical formulations that correlate process behaviour with energy consumption [23, 24].

However, in many cases, the manufacturing process comprises several complex reactions through the production chain, making much more difficult or even impossible to derive dynamic models, based only the physics of the problem [25].

Hence, alternative ways based on the observation of processes main variables, external influencing variables and energy consumption data, started to be widely used to derive consumption approximations.

An important consideration to be taken into account regarding models, and true for all types, is that models are simplified representations of real behaviours or systems and independently of how accurate they are, information is always lost on the modelling phase. Hence, certain factors as time, effort and cost should weight on the modelling method choice. Particularly, deriving mathematical formulations from the analysis of process dynamics may be a very difficult and long process, whereas deriving the same behaviour through an identification method based on the empirical observation may save time and effort. Furthermore, if a model is stored as a collection of signals, the loss of information and modelling effort is transferred to the computational algorithms used to estimate relations between data.

its or and is often used as a replacement of cement.

Another driving factor is that companies often maintain information regarding energy consumption along with usage patterns and process data of their major processes in databases, which combined with the recent rapid evolution of computational algorithms, results in new and better ways to identify causes of energy consumption and determine optimization measures. In fact, many companies have already left traditional statistical[5] and averaging models [6] behind and moved on to regression and time-series or Artificial Intelligence (AI) models [28].

Hence, the next topic of study is a brief overview about these new types of models, describing its structure, advantages, disadvantages and applications,.

2.2.1 Time-Series Regression Models

Models based on regression analysis methods are the most popular in industrial applications, due to their easy of development and ability to predict consumption over long periods of time. [29]

One of the most used regression methods is Support Vector Machine (SVM). SVM was first introduced to solve pattern recognition and regression problems, however rapidly spread into many others fields, including the main topic of this work: energy consumption identification and forecasting [30].

SVM uses only past and present data of observed systems, requiring no foreknowledge of physical properties to forecast energy consumption. The type of data used to build the regression can be either from process main variables and energy consumption or even external influencing variables (e.g. temperature) [31]. In fact, SVM has already been used to forecast energy consumption of a Manhattan skyscraper using only energy consumption data and temperature measurements [32].

In literature, several types of SVM are found, which normally differ on the type of regression. Main derivations are [31]:

- **Linear SVM (L SVM)** — Uses a single linear regression descriptor, thus the prediction is the result of weighted inputs.
- **Non-linear SVM (NL SVM)** — Where the regression function is non-linear;
- **Locally Linear SVM (LL SVM)** —Instead of one global linear descriptor, the input state-space is decomposed into several local linear regressions;
- **Locally Non-linear SVM (LNL SVM)** — Similarly to LL SVM, but instead of one global non-linear descriptor, locally non-linear descriptors are used;

Particularly, LL SVM and LNL SVM are often followed with classification algorithms (e.g. K nearest neighbours) to determine when local regressions should be performed [31].

[5] Traditional statistical methods simply correlate the energy consumption or a certain energy index to production factors [26].

[6] Averaging models are based on linear combinations of consumption from similar days[27].

SVM has been widely used in several applications, such as: forecasting energy consumption of buildings based on weather conditions and season of the year [32, 28] or even occupancy and infrastructure data [29]; Similarly, non-linear dynamic models for HVAC systems have been built using SVM based on temperature and relative humidity [33]. Also, certain complex industrial processes such as furnaces and rotary kilns have been modelled using SVM methods [34].

Regarding another regression method: Recursive Least Squares (RLS) algorithm is commonly used for estimation purposes in a multiple-regression model. Once regression parameters are obtained, a prediction equation can then be derived to predict a continuous output as a linear combination of one or more inputs. This method as been used widely and its popularity may be attributed to the interpretation of model parameters and ease of use. [35, 36].

Another regression method commonly used in industrial applications is the regression tree. In the regression tree method, data collected from observed systems is decomposed into branches and leafs, which are created by recursive partitioning of the input variables until a certain stop criterion is met. Hence, the resultant model is a set of rules that can be interpreted as a tree diagram [28].

Furthermore, regression trees can be used for continuous and categorical variables[7], which is a very important feature given that in energy consumption identification, some causes of consumption may be related to categories (e.g. machine operator A). When categorical variables are used, the regression tree is commonly named as decision tree [28].

A major advantage of regression/decision tree over other regression methods is that it produces a model which represent interpretable rules or logic statements very similar to the usual flowcharts used in the industry.

A great advantage on using regression methods, in general, is that the identification of energy consumption is made based only on the measured data without a deep knowledge of how the processes work, making this method very easily applicable in different types of plants with few model adjustments.

However, the biggest disadvantage is that although the relation between data may be ascertain, there is no mechanism to identify the underlying cause of consumption.

2.2.2 Artificial Intelligence Models

Models based on artificial intelligence have also been largely used for building energy predictions. Particularly, Artificial Neural Networks (ANN) have earned the consensus among researchers as one of the best estimation and prediction methods. Hence, given the impact of this method on the overall research community a more detailed description follows.

[7] A categorical variable, also called nominal variable, is a variables that is only allowed to assume a certain value that belongs to a category with no specific order.

ANN models are inspired by the biological behaviour of nervous systems, particularly the hypothesized process of how the cognitive system of the human brain is able to learn from previous known situations in order to predict results for new situations. Of course, for extrapolating into the future, the human brain, thus also the ANN, has to go through a learning process, often called in ANN as network training [29].

Since it is not well understood how the nervous system works, in particular how a neuron is arranged, several models for ANN were created.

The most commonly used is the *feed-forward* model, in which the neurons are arranged in layers. It is called *feed-forward* (FFANN) as the first layer is the entry point to data from outside of the network, which then streams through the following layers until it reaches the output layer, where the evaluated result is supplied. Typically, there are several neurons on the input layer, one for each variable. However, in the output layer it is normally used only one, since the goal is to learn how multiple input variables are correlated with a single observed result (e.g. how temperature and humidity are related with energy consumption). Between the input and output layers, an ANN may have several layers, often called hidden layers, or not even one. Also, there is no generic rule to specify how many neurons are in each layer, hence the most common method is to train and test the network with several configurations, using the one that has the best prediction result. In the FFANN model, neurons from one layer are only connected to the ones in the next and so on (another reason for the *feed-forward* term) [37, 38]. Figure 2.6 (A) depicts the architecture of an ANN with one input layer with two neurons, 2 hidden layers with 3 neurons and 1 output layer.

Each neuron comprises a combination function that weights the inputs, normally by a linear combination and a transfer function. A typical transfer function used is the Sigmoid function[8] [28]. Figure 2.6 (B) depicts the architecture of a single neuron.

ANN have been widely used for several purposes, namely to forecast energy consumption of commercial [37] and residential buildings [40], using information regarding floor area, air conditioning, building grade, year when it was built, among others. ANN are also used to forecast electrical energy consumption for several industrial processes, using external variables that may impact consumption such as temperature and type of day [41, 42]. In control fields, ANN are used to estimate the behaviour of several dynamic processes, namely furnaces and rotary kilns in cement plants, which comprise complex thermodynamic reactions [43, 44], and for forecasting fuel consumption in rotary kilns as well [45].

A disadvantage of ANN is that it does not provide the parameters that correlate the input variables for testing and interpret their significance. Moreover, the ANN has to be trained with a significant amount of data to provide accurate predictions. Also, before the training stage a preliminary step of feature and configuration selection is needed. Hence, although ANN are very powerful in terms of prediction accuracy, its complexity make it hard to derive conclusions of the underlying cause or relation between the input

[8]For more information regarding the Sigmoid function, consult [39].

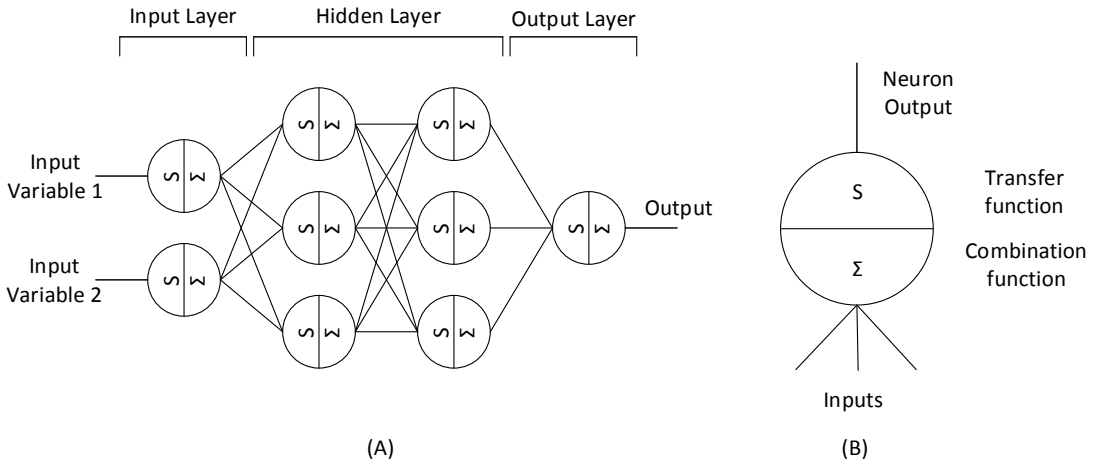

Figure 2.6: (A) FFANN configured with 1 input layer, 2 hidden layers and 1 output layer. Adapted from [37]. (B) Neuron architecture. Adapted from [28].

variables and the output.

Another AI methods that are worth to briefly mention are Genetic Algorithms (GA) and Fuzzy Systems[9] (FS).

Similarly to ANN, GA are also inspired on biological factors. Particularly, the natural evolution process and adaptation of a specific specie to the surrounding environment [46].

On GA a population of a given specie is called group of chromosomes, which evolve and adapt to the environment by exchanging its characteristics via a crossover operation [10], where two parents are selected to produce an offspring. After the crossover operation, chromosomes are subjected to mutation[11]. The participation of each individual chromosome in crossover operations is defined by the evaluation of a selection function, which selects the chromosomes that are best fit to the environment, allowing them to participate more times, whereas the others may be deleted, hence turning possible for the population to evolve. The evaluation of fitness is defined by a fitness function, which takes the chromosome's characteristics/parameters and puts them in a model. Model's estimation is compared with the actual data. The chromosome whose characteristics provide the best estimation is then returned to the environment, and passed to the next generation. [47, 48].

[9] Although some researchers do not consider FS as an AI method as it does not comprise machine learning, the fact that FS are also inspired on human behaviour, particularly on handling with uncertainty, places them on this section.

[10] Crossover operation is the process of evolution from one generation of chromosomes to another, similar to the biological process of reproduction.

[11] The mutation operator acts by changing a chromosome characteristics at a random place with a certain probability.

GA have been used in energy consumption optimization, namely predicting electrical and water consumption in the iron and steel industry [47, 49]. GA and ANN are often combined as in [50], to predict energy consumption in a commercial building or even at larger scale as the energy consumption of an entire country using variables as industrial structure, total population and technological progress [38].

The best advantage of GA is the ability to use accumulative information about the initial unknown environment, which is one of the reasons why it is often combined with ANN. However, a disadvantage is that it is not guaranteed that the algorithm finds the global minimum that best approximates estimation to the real values. Also, in most of the cases parameters convergence tends to be very slow, which is a significant problem for industrial applications, as most of them rely on on-line monitoring systems [48, 49].

Regarding FS, it was first introduced in an attempt to deal with uncertain and ambiguous information, similarly to the behaviour of the human brain and its ability to handle vague information, such as: "low", "high" or "average". Hence, variables go through a so called *fuzzifying* process, in which its numerical values are converted into memberships of certain classes, thus transformed into linguist variables. Afterwards, an inference engine, based on knowledge rules such as: if (premiss) then (conclusion), processes the input variables deriving a conclusion, which is still a linguistic variables that has to be *disfuzzified* into the final output: a numerical value. This concept of *fuzzyfication* and *defuzzification* is in contrast with the traditional binary value "1" or "0" [16]

Vagueness and ambiguity are often found in engineering problems, specially when input/output relations exist (e.g. dynamical systems). Hence, FS have been widely used to tackle those problems in several fields, such as: control systems, e.g. FS are used to control plants or production units [51]; decision making systems [52] and on the area of particular interest of this work: forecasting consumption, where FS are associated with data mining techniques to forecast energy consumption [16].

Similarly to GA, ANN and FS are also often combined as in [53], where the membership functions of FS are continuously adapted according to ANN learning.

A disadvantage of using FS in industrial applications is that for building the inference engine, knowledge-based rules have to be set, hence foreknowledge of processes behaviour is needed.

2.3 Summary

As a summary of the state of the art on existing approaches for energy consumption analysis and optimization, the following conclusions are drawn:

- **Indicator-based** — Practical in industrial environments, however are based on indicators that only reflect the overall consumption of the observed systems, disregarding process level information; Also, the level of detail is relatively low, because it analyses the overall plant consumption and not a specific product line. Hence, it

can not trace back consumption causes to a process level.

- **Knowledge-based** — Allows a fast and general optimization, however requires further knowledge to optimize specific processes, which in industrial plants is a drawback given processes diversity.

- **Resource-based** — Allows an optimization of the energy supply and storage, however disregards consumption of the end-consumer;

- **Simulation-based** — It is able to forecast and identify how energy is used. However current existing models have shown difficulties in explaining certain energy consumption patterns and underlying causes.

Concerning the state of the art energy models:

- **Time-Series and Regression Models** — Allows the identification of energy consumption based only on observed process/system behaviour, consumption and other variables that are considered to have impact on the system, without a deep knowledge of how the processes work. Thus, are applicable in different types of plants with only a few model adjustments. Particularly, regression trees have shown a great advantage on being able to handle categorical and continuous variables, and also for providing a tree look-a-like view, very similar to the wide used flowchart. Regarding RLS algorithm, it is a very practical and efficient method for on-line estimation and energy consumption prediction. However, the biggest disadvantage, and true for the time-series and regression models addressed, is that although the relation between data may be ascertain, there is no structural way to differentiate between inputs/outputs and external factors that may cause consumption, thus making difficult to identify the underlying consumption cause.

- **Artificial Intelligence Models** — Allows a very accurate estimation of energy consumption and is able to consider several variables in the estimation process. However there are several disadvantages on the methods addressed. In particular, ANN do not provide parameters used for prediction in a form that can be interpreted or analysed in terms of significance, thus although a very accurate estimation is performed, the causes of consumption are still unknown. Also, the ANN have to be trained first with a significant amount of data samples in order to produce reasonable estimations. Regarding GA, it is a very good method to estimate parameters when initial information about the process is unknown, however the algorithm convergence tends to be extremely slow, which is a drawback in on-line energy consumption estimation; Lastly, FS is a suitable method to deal with uncertainty, however foreknowledge about the process is needed to define the best set of inference rules.

As a driving factor for the proposed approached of this dissertation, the existing methods and models do not provide a structural way to incorporate influences or external factors on the energy consumption estimation, thus not being able to identify the causes of consumption or to explain certain energy consumption patterns. Hence, this dissertation proposes an approach to tackle this existing gap.

2.4 Proposed Approach

The proposed approach of this dissertation is based on the identification of energy consumption by correlating time-based energy readings with the underlying operating context of a given plant, process or production unit.

The operating context is defined by the influence of external variables on plant operation. Hence, the mapping between operating context and consumption patterns allows the identification of the cause-effect relation. As external variables may be either continuous (e.g. temperature) or categorical (e.g. human operator A), regression trees are particularly suitable for context identification in industrial applications, as this method is able to handle both types. Furthermore, the tree look-a-like representation of external factors which impact consumption is very useful for identifying when and why certain consumption patterns are observed, leading to a better design and implementation of optimization measures. As the decision-making process is shared between process engineers and management teams, tree diagrams (due to the similarity with flowcharts) provide a natural support for understanding and explaining to teams that are not accustomed with algorithms, the causes of consumption. The context identification method is described on the next chapter - Context Identification.

As there are several different processes on industrial applications, a method for estimating energy consumption that does not rely on deep knowledge about the process behaviour is needed. Hence, an on-line multi-model approach using multiple RLS algorithms (one for each operating context) is used to estimate energy consumption while the plant is operating. Each RLS algorithm estimates the energy consumption for a specific operating context, defined by the output (identified context) of the regression tree algorithm, giving rise to the Context-based Energy Models (CBEM) concept. By using multiple RLS algorithms, the modelling and prediction of energy consumption based on context identification, can be carried out dynamically, which enriches the applicability of the proposed approach in the fields of energy savings detection and optimization on industrial plants. How regression trees and multi-models play together, to provide a context-based energy consumption identification is described on chapter 3 - Context-based Energy Models.

3

Context Identification

This chapter comprises an approach for modelling external factors (introducing the concept of context variable) and depicts how they impact plants operation. An automatic context identification algorithm, based on regression tree is described, along with computational simulations for theoretical demonstration of the concepts involved.

3.1 Modelling Context Influence

3.1.1 Context Variables

In industrial applications the term plant is often used to define a system with certain inputs and outputs that performs a task. This wide definition, mostly used by process engineers, allows a single process (for what the term is most commonly used), a production unit, the whole factory or in contrast a tiny valve to be represented (regardless of the complexity) in a structured way that puts into perspective how the plant transforms inputs into outputs, e.g. in manufacturing processes, an input is normally a flow of raw material, in t/h, and the output also in t/h is the result of chemical, physical or mechanical transformations. This conceptual representation is very suitable for control or production purposes as it puts together the main variables to be taken into consideration for the design of production or control system where the understanding of how plant behaves (how it transforms inputs into outputs) is fundamental, in order to keep it under a desirable operating state.

However, from an energy consumption point of view, the main focus is to understand the way of how energy is used while the plant is operating.

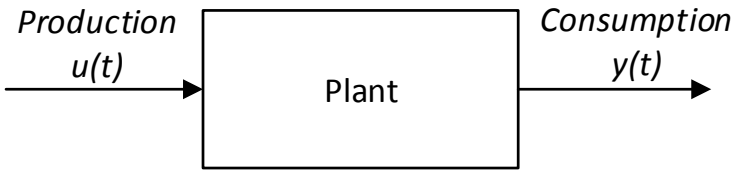

Figure 3.1: Representation of a plant from an energy consumption point of view.

Hence, for production or control purposes a plant represents a system that performs a transformation of an input into an output, whereas for energy consumption purposes a plant is represented as a system that consumes energy for performing the transformation. For keeping the conceptual representation of a plant as a system without the loss of generality, consider that from an energy consumption point of view a plant is a system that takes as input, variables related to production (e.g. production rate) and has an output the energy/power required for production (e.g. watt, specific consumption) as depicted in figure 3.1. Following this research line, and to put together production and consumption variables into a structure, a plant can be defined by the following equation:

$$y(t) = f\{u(t)\} + e(t) \tag{3.1}$$

Where y is the plant's consumption; u is a production variable; f denote a mapping function between production and consumption and e is external factors (not related to production, hence not defined by u) that may impact consumption.

However, as depicted on the previous chapter, state of the art energy models have shown difficulties in explaining certain energy consumption patterns based only on the relation between production and consumption. This is mostly due to the fact that, particularly in industrial environments, the plant is strongly exposed to external factors, namely machinery factors(e.g. equipment condition), human factors (e.g. operators degree of expertise), ambient factors (e.g. temperature, humidity), among others.

If equation 3.1 would be interpreted from a production or control point of view, e would be considered as an external factor that undesirable impacts plant's operation and is not under direct control for maintaining the plant under a desirable operating state, often called a disturbance. In a sense, some external factors (e.g. temperature) can be considered as disturbances, however from an energy consumption point of view these external factors hold significant information regarding the context in which the plant is operating. Hence, e comprises two different types of variables: context variables and disturbances.

One way to distinguish between them is to analyse how both impact consumption (from a production or control point of view, their impact is the same, as both affect the

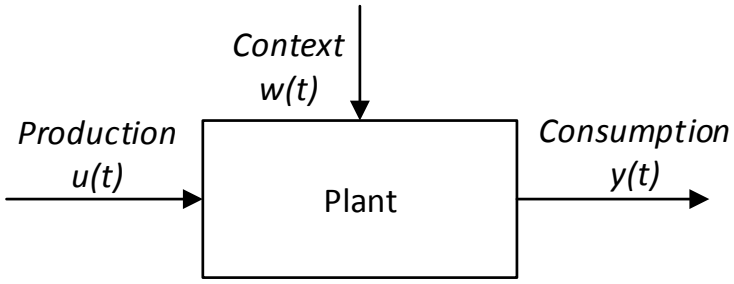

Figure 3.2: Representation of a plant from an energy consumption point of view, under the influence of context.

plant operating state). Starting from disturbances, although they impact plant's operation, the consumption observed is due to the plant's reaction to the disturbance, which means that disturbances occur on a similar time-scales as plant's time constants, and after they end, the plant is able to return to its original state.

On the other hand, context variables define a context in which the plant is operating (lets call it operating context), which means that they impact consumption on time-scales that goes beyond plant's time constants. In this way, a plant does not only react to a context variable, but its behaviour (the way inputs are transformed into outputs) is influenced by it, which is reflected in certain energy consumption patterns.

Furthermore, the definition of operating context, provides a natural support to simultaneously correlate energy consumption with production and context variables. Hence, a suitable interpretation is that consumption is not only related to production but also the context in which production takes place as depict in figure 3.2. Hence, the mapping function between production and consumption is context dependent.

Therefore equation 3.1 has to be re-written as follows:

$$y(t) = f\left\{u(t), w(t)\right\} + e(t) \tag{3.2}$$

Where w are context variables. According to the this approach e is a disturbance on plant's operation that cannot be explain in terms of energy consumption by production or context factors. Also e may comprise consumption patterns (expected to be almost insignificant) from other influences (either derived from context or not) that are not taken into consideration on the mapping function.

In order to summarize the variables at stake, it is proposed the following description:

- **Production variables** — Variables related to plant's operation;
- **Consumption variables** — Variables that depict consumption or performance indicators;

- **Context variables** — Variables that define an operating context and which impact plant's operation for periods of time that goes beyond plants time constants. By this definition, although brisk changes are allowed, a minimum period of steadiness must be followed;

- **Disturbances** — External variables that have impact on plant's operation, in the same time-range as of the process time constants, whose consumption cannot be explained either by production or context factors.

Following this research line, the next step is to determine how context variables define the operating context, driving this study to the next section.

3.1.2 Operating Context, Regions and Centroids

Context variables, similarly to any other time-varying variables, change as time goes by. For instance, temperature may have brisk changes at sunrise and sunset, but tends to become spatially uniform around a nearly constant value between those boundaries of abrupt change, hence is in accordance with the definition of context variables introduced in the previous section, which states that although brisk changes are allowed, a minimum period of steadiness must be followed. Another example is a context variable that identifies which operator is controlling a plant on a certain period (in this case the context variable is categorical). When the working shift ends, operators switch (which is a binary commutation), thus the operating context changes. In this case the minimum period of steadiness is defined by the duration of a working shift, whereas in case of temperature it may be defined by season of the year and/or time of the day (morning, afternoon and night).

Hence, the period of steadiness vary from one context variable to another. Furthermore, it may not even be uniformly spaced.

Hence, for maintaining the degree of generalization, a suitable approach is to state that an operating context at a given time, is identified by the nearly constant value of a context variable (lets call it centroid), between the boundaries (lets call it region) of the period of steadiness. As time goes by, obviously the operating context changes.

By this approach, if a model is to be estimated to identify energy consumption based on context variables, it is possible to track down in which operating context (represented by the centroid) a certain energy consumption pattern/footprint was observed. Hence, causes (context) are mapped into effects (consumption), providing a structured way to correlate production, context and consumption.

To summarize the concepts involved, consider the following description:

- **Context region** — A period of steadiness where the context variable is kept nearly constant;

- **Centroid** — The mean value of the context variable over a context region;

- **Operating context** — Plant operation when influenced by a certain context, identified by a centroid.

Henceforward consider the nomenclature to define a context region as R, and in the case of a specific region n belonging to a specific context variable w, the nomenclature is R_n^w. If the context variable is implicit, then a nomenclature as R_n may be used. Similarly, a centroid is represented as c, and in the case of a centroid in a specific region, the nomenclature is c_{R_n}.

For better depicting what regions, centroids and operating context represent, consider a simple simulation example described in the following section.

3.1.3 Simulation Example

The purpose of this simulation is to illustrate de context identification concepts, applied to an industrial plant, e.g. a milling unit, as depicted in figure 3.3.

In industrial milling units the different mixture of materials to be ground define the type of product under production. Depending on the mixture, the resulted blended material is more or less hard, which may require the re-circulation of material within the milling circuit in order to achieve a desirable level of narrowness, consuming more energy in this process.

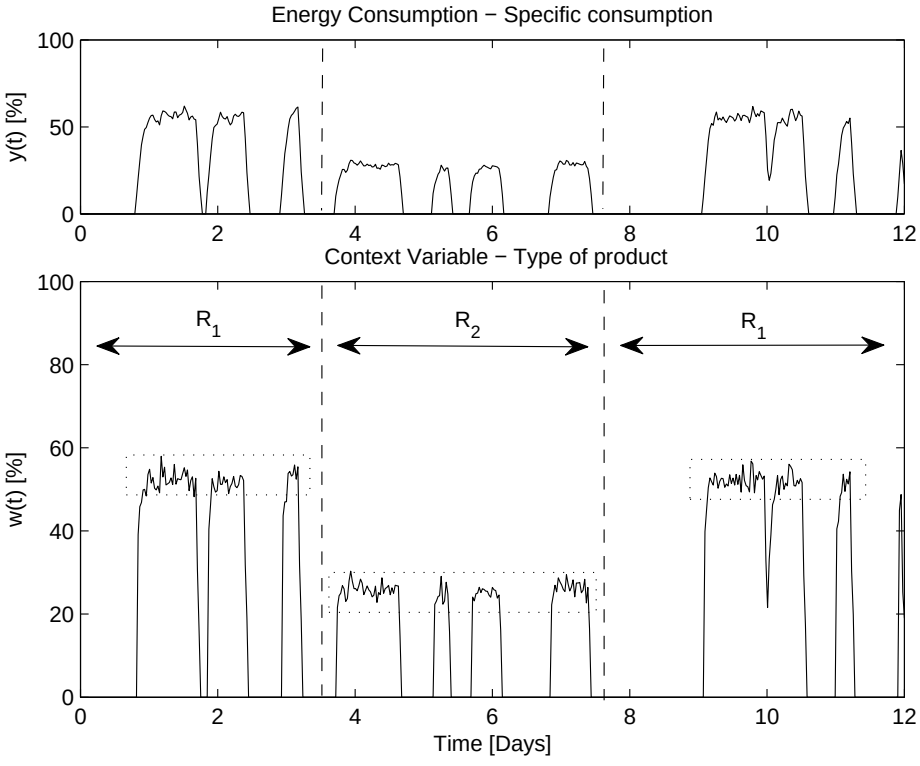

Figure 3.3: Context influence on energy consumption. The case of a milling unit.

Therefore, in this case, the type of product being milled may be considered as a context variable. The flow of ground material a production variable and as a consumption measurement, consider the mill specific consumption[1].

During the mill production time, two different types of product were produced. Product type 1, represented by an approximate constant value of 50%, was milled during the first 3 days, whereas product type 2, represented by approximately 25%, was milled between the end of the 3$^{\text{th}}$ and the end of the 7$^{\text{th}}$ day. Afterwards, product type 1 was under production again. Context region R_1 defines the operating context when product 1 is under production, and context region R_2 for product type 2.

As it is depicted specific consumption shows quite different patterns when the operating context changes. This is due to the type of product being produced. As the mixture of materials needed to mill product type 1 is harder that the mixture for product type 2, more energy is required for the milling process per unit of mass, as materials need to be re-circulated.

In order to carry out dynamically (while the plant is operating) the context and consumption identification, it is required a computational method able to automatically identify context regions, by finding its boundaries and calculate the average value of the context variable within the region, thus the centroid. Which drives this work to the next step: automatic context identification.

3.2 Automatic Context Identification

A way to define regions is by successively sub-divide, or partition, the context variable space into smaller groups until the resultant chunks are nearly steady constant over an average value or a certain period of steadiness is guaranteed. This method is called recursive partitioning. A well known and studied algorithm - Regression Trees (RT) is a suitable method for the partition of both continuous or categorical variable, which represents the partition on a tree diagram.

3.2.1 Regression Tree Algorithm

The issue of finding structural breaks can be briefly described as follows [54, 55].

Consider that a context variable $w(t)$ (either categorical or not), is characterized by L context regions with $L-1$ breaks, in which $w(t)$ can be approximated by a mean value. Hence, the partition of $w(t)$ is defined as:

$$w(t) = c_{R_l} + \xi(t), \quad l = 1, \cdots, L \tag{3.3}$$

[1]Signals are normalized, as this simulation is only for demonstration purpose.

Where c_l is the centroid's region, ξ is the error term of the partition and t is the partition of the time vector into regions as follows:

$$t = T_{l-1} + 1 \cdots T_l, \quad l = 1, \cdots, L \tag{3.4}$$

Assuming the convention $T_0 = 0$ and T_L is the series length.

The problem is to identify the set of unknown commutation points (i.e. boundaries) that best partition the context variable:

$$P(L) = \{(1, \cdots .T_1), \cdots, (T_{l-1} + 1, \cdots, T_l), \cdots, (T_{L-1} + 1, \cdots, T_L)\} \tag{3.5}$$

Hence, $P(L)$ defines the partition of w into L regions:

$$P(L) = \{R_1, \cdots, R_l, \cdots, R_L\} \tag{3.6}$$

Where a region R_l is given by:

$$R_l = (T_{l-1} + 1 \cdots T_l) \tag{3.7}$$

One way to solve the problem is to use the least squares principle, so that the estimated regions:

$$\hat{P}(L) = (\hat{R}_1, \cdots, \hat{R}_l, \cdots, \hat{R}_L) \tag{3.8}$$

are such that:

$$\hat{P}(L) = argmin_{\hat{P}(L)} SSR(P(L)) \tag{3.9}$$

Where the objective function $SSR(P(L))$, is the sum of squares residual (SSR), defined by:

$$SSR(P(L)) = \sum_{l=1}^{L} \sum_{t \in R_l} (w(t) - c_{R_l})^2 \tag{3.10}$$

And c_l is the mean value (centroid) of $w(t)$ over a region R_l.

$$c_{R_l} = \sum_{t \in R_l} \frac{w(t)}{T_l - (T_{l-1} + 1)} \tag{3.11}$$

A Least Squares Regression Tree (LSRT) algorithm presented in [54, 56], depicts a method to minimize 3.10. By this method, a node is split into its left and right descendants in order to reduce the deviance of the dependent variable, fitting to each descendant the mean of the dependent variable. The selected split is the one that minimizes the between-group SSR. Once the partition of a node is performed, the splitting process is recursively applied to each descendent.

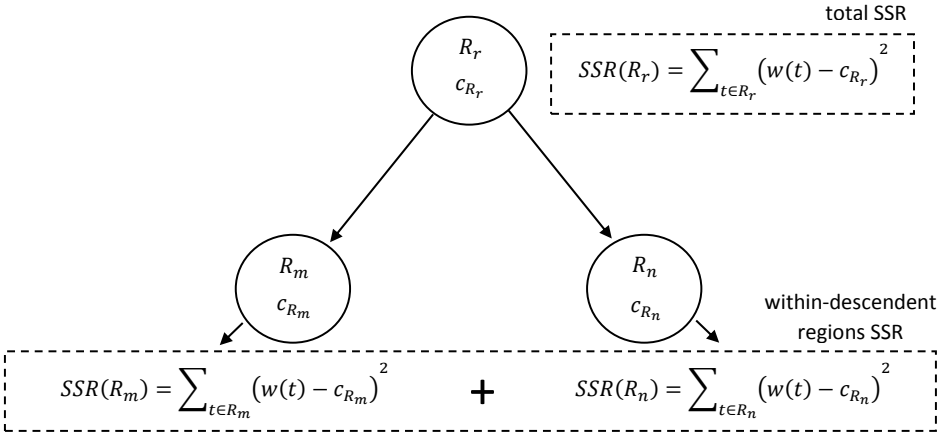

Figure 3.4: Partition criterion of a region R_r into R_m and R_n. Adapted from [55].

By the same principle w is recursively partitioned, by finding the best set of regions which minimize the SSR between w and each c. For instance, at a given point of the partition process, region R_r is split into two descendent regions R_m and R_n as depicted in figure 3.4.

The LSRT split criterion is the selection of the descendent regions which minimize:

$$SSR(R_r) - (SSR(R_m) + SSR(R_n)) \tag{3.12}$$

Where $SSR(R_m)$ and $SSR(R_n)$ are, respectively the SSR of the left (R_m) and right (R_n) descendent regions. $SSR(R_r)$ is the SSR of the parent region R_r:

$$SSR(R_r) = \sum_{t \in R_r} (w(t) - c_{R_r})^2 \tag{3.13}$$

As R_m and R_n are an exhaustive partition of R_r, $SSR(R_r)$ represents the total sum of squares residuals at region R_r and $(SSR(R_m) + SSR(R_n))$ the within-descendent regions sum of squares residuals.

Thus, the splitting criterion 3.12 consists in minimizing:

$$SSR(R_m) + SSR(R_n) = \sum_{g \in \{m,n\}} \sum_{t \in R_g} \left(w(t) - c_{R_g}\right)^2 \tag{3.14}$$

Therefore, the LSRT split criterion 3.14 corresponds to the objective function 3.10.

Partition recursively continues, hence creating more regions, until a certain stop criterion is achieved. There are several stop criterion methods. However, from the definition of context variable presented before, a minimum period of steadiness (lets call it τ) has to be followed although brisk changes are allowed. Therefore, the stop criterion is to stop

partition when a region R_l is less than the τ:

$$R_l < \tau \tag{3.15}$$

Another stop criterion that may play along with 3.15, is to stop the partition when the centroid calculated gives less than some minimal amount of extra information. This can be achieved by setting a tolerance value κ, proportional to the SSR per region. When the quadric error drops below the threshold ε, partition is stopped:

$$\varepsilon < \kappa \cdot \sum_{t \in R_l} (w(t) - c_{R_l})^2 \tag{3.16}$$

If a prior knowledge of how the context variable changes over time is known, it can be used to initialize τ and κ, allowing a better context region partition and a more accurate centroid calculation. The tolerance value κ is normally initialised with a very small value, e.g. 10^-6 in order to give the best approximation as possible. However, lets maintain κ has it gives more liberty degrees on the algorithm's tuning, although most of the times it will be initialized with the default value.

Consider now the case where w is a categorical variable. If w is categorical, then the best set of regions $\hat{P}(L)$, obtained by performing LSRT with the split criterion defined in 3.14, turn $SSR(P(L)) = 0$. Thus, $\xi = 0$. Which means that from 3.3, the context variable is fully partition into $w(t) = c_{R_l}, l = 1, \cdots, L$ where t is given by 3.4.

Indeed, LSRT provide a practical and elegant way for identifying context. The partition is represented as a tree whose several paths (leafs) to the terminal node provide the context region boundaries and the terminal node provides the centroid, which only applies on that region.

For a better understanding of the LSRT algorithm, consider the example depicted in the next section.

3.2.2 Simulation Example

This simulation example depicts the LSRT algorithm applied to a context variable as depicted on figure 3.5.

The context variable is approximately piecewise constant over 3 regions. Ideally, the first region R_1, should occur from the beginning of the simulation until nearly 12$^{\text{th}}$ minute and has a centroid value of approximately 80%. After minute 12 until the 28$^{\text{th}}$ minute, the second context region R_2 is defined, with a centroid of about 40%. After minute 28 and until the end of the simulation, is the delimitation of the third context region R_3 with a centroid value of about 80%.

In order to automatically identify the context regions and calculate the respective centroids, a regression tree algorithm is initialized with a minimum period of steadiness $\tau = 10$ that guarantees that context regions should not be shorter than 10 minutes. The centroid tolerance was initialized with the default value of $\kappa = 10^{-6}$.

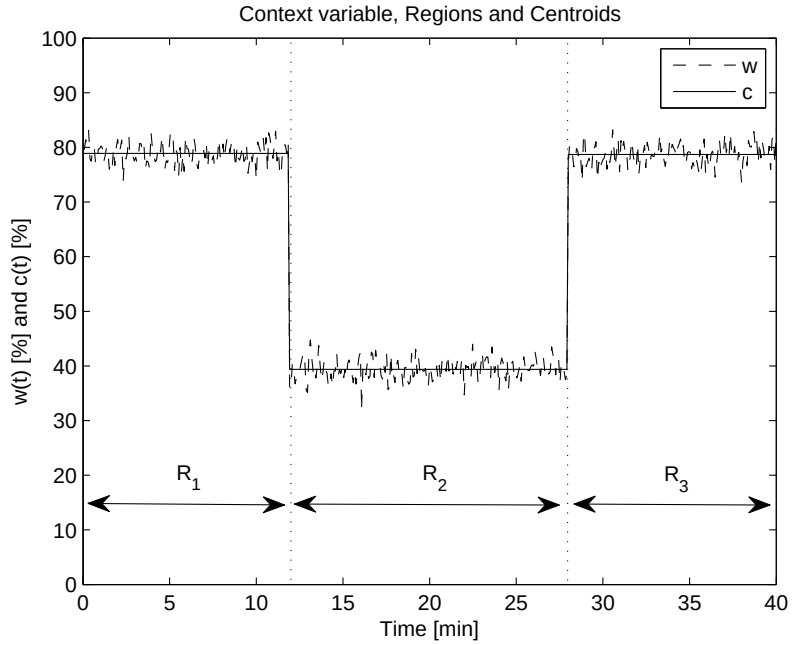

Figure 3.5: Context variable, centroid calculation and region boundaries delimitation

Figures 3.5 and 3.6 depict, respectively the result of LSRT in the form of time-series and tree diagram. Table 3.1 depicts the exact results. As shown, although the context variable presents brisk changes, LSRT is able to identify boundaries and calculate centroids on the regions where w ranges around nearly a constant value.

However, the use of numerical algorithms to automatically identify context raises a

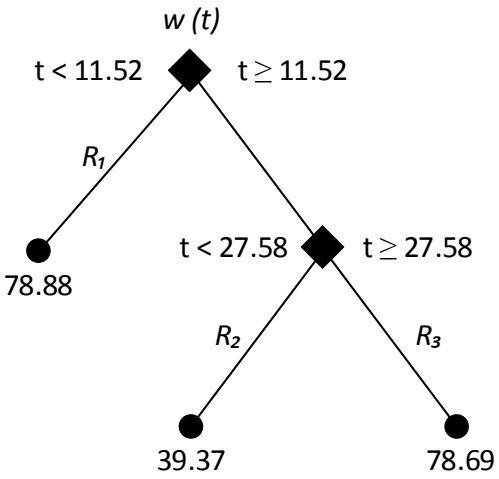

Figure 3.6: Tree diagram of the LSRT applied to w.

Table 3.1: Regions, Centroids and Region Boundaries results.

Region	Centroid [%]	Region Boundaries [min]
R_1	$c_{R_1} = 78.88$	$0 \leq t < 11.52$
R_2	$c_{R_2} = 39.37$	$11.52 \leq t < 27.58$
R_3	$c_{R_3} = 78.69$	$27.58 \leq t < 40$

new conceptual problem, particularly emphasize on this simulation. Consider the centroids of the regions R_1 and R_3. As it is depicted, they are very similar. This can be interpreted in two ways:

1. Regions R_1 and R_3 represent the same operating context. Therefore the LSRT algorithm has to be adapted, bringing complexity and consequences (e.g. increased partition error ξ and entropy on finding the best splits).

2. As consumption models will be estimated according to operating context, the choice of selecting a single model to estimate consumption on both R_1 and R_3 regions or two distinct models for each one, should be passed to a selection function (commonly used on multi-model architectures).

The 2nd interpretation is the most reliable. This way, the selection function defines (e.g. setting a validity range for each model) if different models should be learned on each region or not. Furthermore, different estimation methods (or other systems) are able to use LSRT outputs and interpret the information according to it each one purpose. This work-line will be described further on this dissertation.

LSRT shows to be a promising way for automatically identifying context. However, what if when a plant is operating, it is exposed to multiple context variables? And how would LSRT be used to automatically identify the operating context? These questions are answered in the next section.

3.3 Multiple Context Variables

In industrial environments, a realistic approach is to consider that a plant may be exposed to multiple external factors at the same time, which all together influence plant's operation, hence influencing energy consumption. Therefore, in order to build more accurate consumption models and to completely identify an operating context, the influence on plant's operation and consumption, should be interpreted as a result of all relevant context variables acting independently, although simultaneously on plant's operation.

3.3.1 Operating Context in the Presence of Multiple Context Variables

The LSRT method described in the previous section, showed a way to decompose a context variable into a set of regions, where for each, a centroid is calculated. However,

when several context variables are at stake, an operating context is not completely defined by one single context variable, but instead, by the intersection of all the context variables that are affecting plant's operation at a given time.

Hence, consider that two different context variables w_1 and w_2 are influencing plant's operation.

As described previously, context variable w_1 is partition into L regions:

$$P^{w_1}(L) = \{R_1, \cdots, R_l, \cdots, R_L\} \tag{3.17}$$

And context variable w_2 into M regions:

$$P^{w_2}(M) = \{R_1, \cdots, R_m, \cdots, R_M\} \tag{3.18}$$

Where P^{w_1} and P^{w_2} are respectively, the partition of w_1 and w_2.

Hence, if the operating context is to be defined as the intersection of w_1 and w_2, the operating context is given by:

$$P^{w_1}(L) \cap P^{w_2}(M) \tag{3.19}$$

Which results on context regions of the following form:

$$R_{l,m}^{w_1,w_2} = R_l^{w_1} \cap R_m^{w_2} \tag{3.20}$$

Thus, in order to identify the context region $R_{l,m}^{w_1,w_2}$, centroids of each individual region have to be joined:

$$c_{R_{l,m}^{w_1,w_2}} = c_{R_l^{w_1}} \cup c_{R_m^{w_2}} \tag{3.21}$$

By this interpretation, LSRT method still applies. The LSRT is applied to each context variable individually, defining context regions centroids for each one. After this first iteration, context regions that overlap are intersected and centroids are joined, in order to uniquely identify the operating context.

Consider the following simulation example for a better demonstration of the approach proposed.

3.3.2 Simulation Example

In this simulation example, 2 different context variables are simultaneously influencing plant's operation. The first iteration is to separately apply a regression tree analysis for each context variable in order to calculate centroids and region as depicted in figure 3.7. The context variable w_1, is the same as the previous simulation example and has 3 context regions. Context variable w_2 has 2 context regions, one approximately kept steady at 80% and the other at 30%, switching nearly on the middle of the simulation at minute 20.

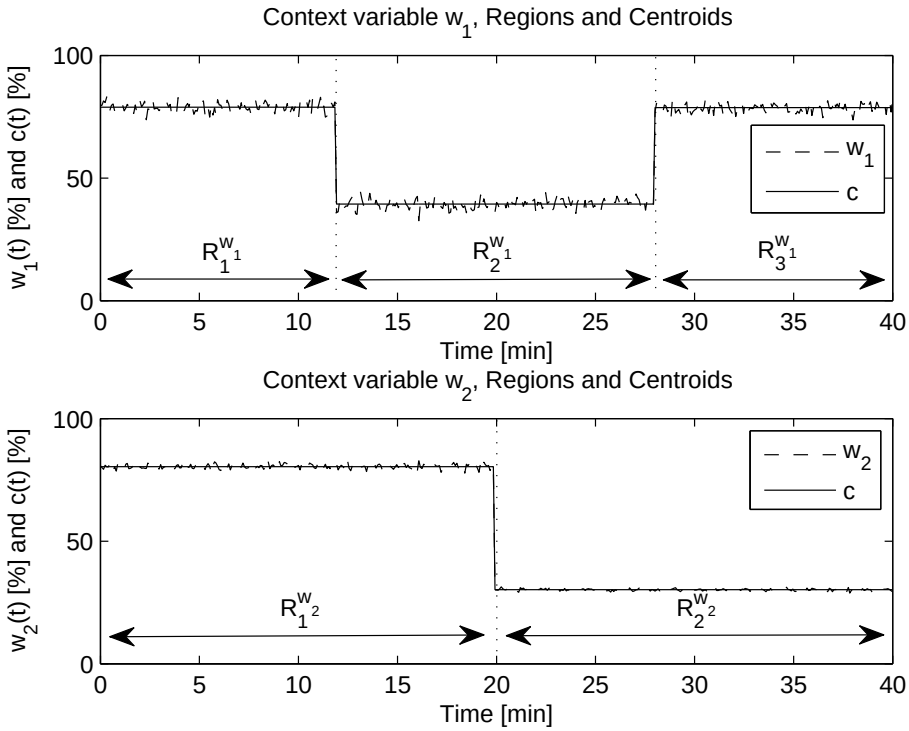

Figure 3.7: Context variable w_1 and w_2, regions identification and centroid calculation

Both regression tree algorithms were initialized with a minimum period of steadiness of $\tau = 10$ minutes and a centroid tolerance of $\kappa = 10^{-6}$. Although τ and κ were initialized with the same values for both regression algorithms, each one can be initialized with different values, allowing a better tuning for each context variable. Figure 3.8 illustrates the tree diagram for w_1, on the left side, and w_2, on the right side.

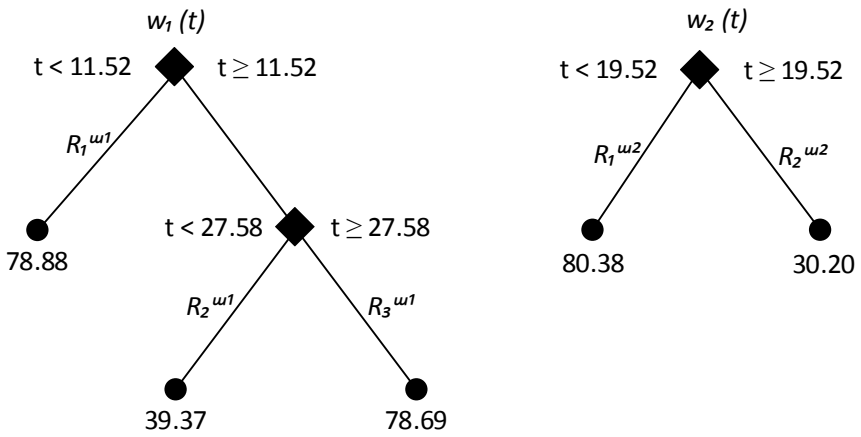

Figure 3.8: Tree representation for context variable w_1 and w_2.

Table 3.2: Regions, Centroids and Region Boundaries for w_1.

Region	Centroid [%]	Region Boundaries [min]
$R_1^{w_1}$	$c_{R_1^{w_1}} = 78.88$	$0 \leq t < 11.52$
$R_2^{w_1}$	$c_{R_1^{w_1}} = 39.37$	$11.52 \leq t < 27.58$
$R_3^{w_1}$	$c_{R_3^{w_1}} = 78.69$	$27.58 \leq t < 40$

Table 3.3: Regions, Centroids and Region Boundaries for w_2.

Region	Centroid [%]	Region Boundaries [min]
$R_1^{w_2}$	$c_{R_1^{w_2}} = 80.38$	$0 \leq t < 19.52$
$R_2^{w_2}$	$c_{R_1^{w_2}} = 30.20$	$19.52 \leq t < 40$

Tables 3.2 and 3.3 depict the numerical regression analysis results for w_1 and w_2, respectively.

The second iteration is the intersection of the context regions that overlap, and the union of the respective centroids. Figure 3.9 depicts on the same graph both context variables, so that the regions which overlap are best noticed.

The first intersection occurs from the beginning of the simulation until nearly the 12^{th} minute, in which context region $R_1^{w_1}$ from context variable w_1 is overlapped with context region $R_1^{w_2}$ from context variable w_2. The result of the intersection is a new region $R_{1,1}^{w_1,w_2}$,

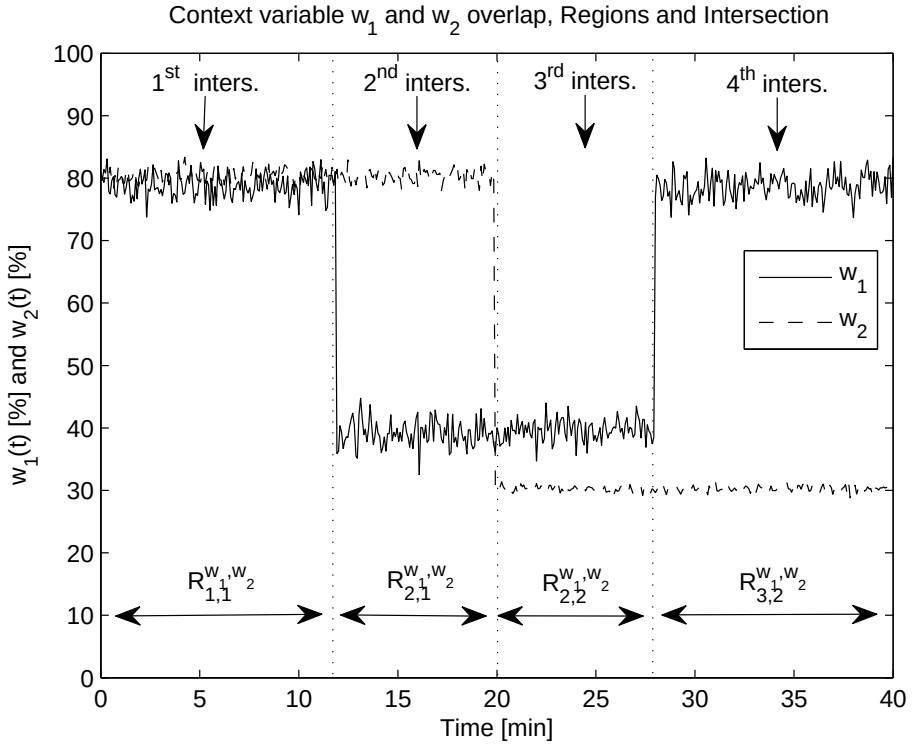

Figure 3.9: Context variables overlapped and region intersections.

Table 3.4: Regions and Centroids intersection.

Region	Intersected Regions	Centroid [%]	Region Boundaries [min]
$R_{1,1}^{w_1,w_2}$	$R_1^{w_1} \cap R_1^{w_2}$	$c_{R_{1,1}^{w_1,w_2}} = \{78.88; 80.38\}$	$0 \leq t < 11.52$
$R_{2,1}^{w_1,w_2}$	$R_2^{w_1} \cap R_1^{w_2}$	$c_{R_{2,1}^{w_1,w_2}} = \{39.37; 80.38\}$	$11.52 \leq t < 19.52$
$R_{2,2}^{w_1,w_2}$	$R_2^{w_1} \cap R_2^{w_2}$	$c_{R_{2,2}^{w_1,w_2}} = \{39.37; 30.20\}$	$19.52 \leq t < 27.58$
$R_{3,2}^{w_1,w_2}$	$R_3^{w_1} \cap R_2^{w_2}$	$c_{R_{1,2}^{w_1,w_2}} = \{70.60; 30.20\}$	$27.58 \leq t < 40$

whose boundaries are from 0 to 12 minutes. The respective centroid $c_{R_{1,1}^{w_1,w_2}}$ is given by the union of both context region centroids, $c_{R_1^{w_1}} = 78.88$ and $c_{R_1^{w_2}} = 80.38$. Thus, centroid value is $c_{R_{1,1}^{w_1,w_2}} = \{78.88; 80.38\}$. Similarly, the second intersection is between context regions $R_2^{w_1}$ and $R_1^{w_2}$ resulting on $R_{2,1}^{w_1,w_2}$. Same logic is applied for the remaining regions.

Table 3.4 shows the numerical results of region intersections, centroid unions and respective region boundaries.

Lastly, the final tree of intersected regions, is given by the intersection of each context variable tree. Tree leafs are intersected, as they represent context regions and tree terminal nodes are joined as they represent centroids, as depicted in figure 3.10.

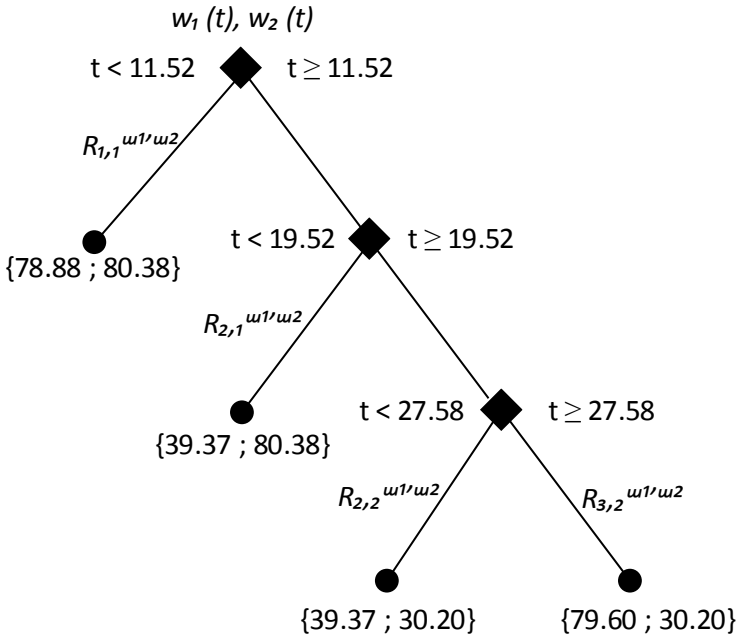

Figure 3.10: Tree diagram of multiple context variables overlapped and region intersections.

Although this approach allows the handling of multiple context variables, if a considerable amount of context variables are used, the

As a summary of this chapter, using LSRT for partition of context variables into operating contexts and identifying each one by a single unique centroid, allows the correlation of context (cause) to a certain consumption pattern (effect). Hence, the estimation of an energy model within the boundaries of a context region, will capture the influence that the context has on consumption, which drives this dissertation to the next chapter: Context-based Energy Models (CBEM).

4

Context-based Energy Models

In this chapter it is presented a method for building context-based energy models. Based on context variables partition, a bank of local models are estimated, in which it is captured the influence of context on energy consumption. A structure for the bank of models is also described.

Several computational simulations are performed in order to demonstrate the concepts involved and for theoretical validation. At the end of the chapter, a comparison with the traditional estimation method is performed to show the contribution of the proposed approach.

4.1 Context-based Multi-model Approach

As described on chapter 2, in industrial applications it is not common to estimate consumption models purely based on the mathematical analysis of plant's physical processes and how they impact energy consumption, due to the complexity of the processes involved and the amount of effort and time required to build them.

Typically, consumption models are built based on a sequence of samples retrieved from production and energy consumption observations when the plant is operating. It was shown and proven [57] that a sequence of sampled observations can be used to approximate the mapping function f, responsible for mapping between production data and energy consumption patterns, presented previously in equation 3.1, and copied again as follows for a better understanding:

$$y(t) = f\{u(t)\} + e(t) \tag{3.1}$$

The mapping function approximation $\hat{f}$ is given by a set of parameters Θ which estimates the relation between the observed present and past measurements of u and y:

$$\hat{f}(\Theta, \varphi(t_k))) = \Theta \cdot \varphi(t_k) \tag{4.1}$$

Where t_k is the discrete time variable, given by $t_k = t \cdot T_s$, where T_s is the sampling period; φ is often called regressor and is a vector of past and present measurements of production and energy consumption data, which is extended as follows:

$$\varphi(t_k) = [y(t_k - 1)\ ;\ \cdots\ ;\ y(t_k - n_y)\ ;\ u(t_k - 1)\ ;\ \cdots\ ;\ u(t_k - n_u - 1)]^T \tag{4.2}$$

Where n_y and n_u denote, respectively y and u order of regression .

By substituting f by $\hat{f}$ on equation 3.1, the result is an estimated energy consumption model, given by:

$$\hat{y}(t_k) = \Theta \cdot \varphi(t_k) + \epsilon(t_k) \tag{4.3}$$

Where $\hat{y}$ is the estimation of y and ϵ is the estimation error.

However, the approach of estimating energy consumption models based uniquely on production and energy consumption data, has shown difficulties in explaining certain consumption patterns, due to disregarding the influence of context variables on the approximation and also for using a unique and global approximation function for the entire plant operation, without taking into consideration that different operating contexts have impact on plant's consumption. In order to distinguish between the current existing method and the proposed approach on this dissertation, lets name the consumption model defined in equation 4.3 as the traditional method.

The approach proposed in this dissertation tries to fill in the gap by taking into consideration that context variables should be included on the mapping function, as they have impact on plants consumption.

By combining production, consumption and context data, the mapping function is enhanced, as it was depicted in equation 3.2 and it is possible to track down the causes of consumption into observed consumption patterns. Hence, from this approach, as context variables affect plant's behaviour, they should be incorporated in the estimation of Θ, which turns it context dependent. Taking this into consideration, the mapping function approximation defined on equation 4.1 is re-written as follows:

$$\hat{f}(\Theta(w(t_k)), \varphi(t_k)) = \Theta(w(t_k)) \cdot \varphi(t_k) \tag{4.4}$$

Furthermore, the partition of context variables into regions, from which operating contexts are identified, provides a natural support for a multi-model architecture. Instead of using a single $\hat{f}$ as the traditional method, the context-based approach proposes that the overall energy consumption estimation should be decomposed into a family of local

maps $\hat{f_M}$[1], one for each operating context:

$$\hat{f}(\Theta(w(t_k)), \varphi(t_k)) = \bigcup_{i=1,\cdots,L} \hat{f}_{M_i}(\Theta_{M_i}, \varphi(t_k)) \tag{4.5}$$

Where L denotes the number of operating contexts and Θ_M the vector of parameters estimated for each one. Each individual map captures the influence of context on energy consumption, by locally estimating Θ_M based on the regression vector valid within the boundaries of a context region. It is assumed that the partitioned regions should be fine enough so that a local model can be estimated. As a context region should be kept steady for a minimum period of steadiness, it is guaranteed that a minimum number of samples are used for estimation.

Thus, the context-based energy consumption model is:

$$\hat{y}(t_k) = \bigcup_{i=1,\cdots,L} \hat{y}_{M_i}(t_k) \tag{4.6}$$

Where $\hat{y_M}$ is a local model:

$$\hat{y}_M(t_k) = \Theta_M \cdot \varphi(t_k) + \epsilon(t_k) \tag{4.7}$$

By this way, the complicated mapping function is simplified by the partition of context variables into operating contexts and using local models to estimate consumption on each one. Figure 4.1 illustrates the proposed approach.

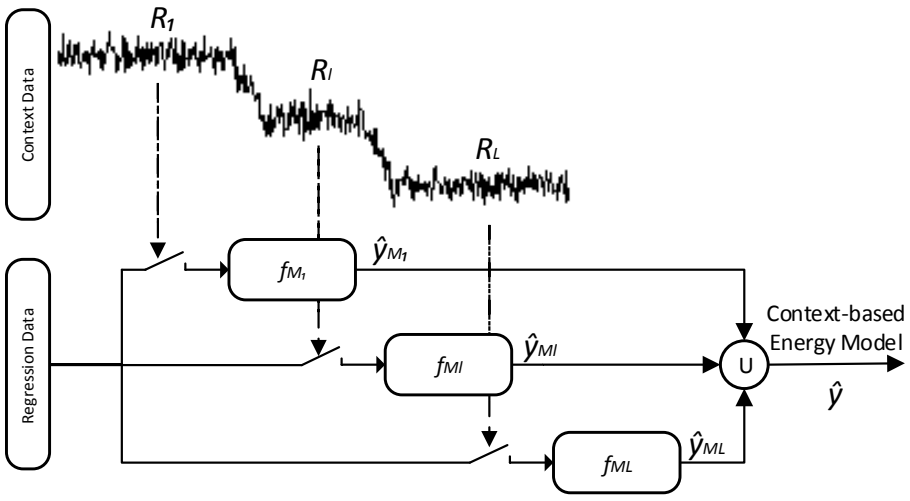

Figure 4.1: Context-based multi-model approach.

[1]Note that the nomenclature $\hat{f}_M$ is used because a model may estimate consumption on more than one region, as it will be depicted further on this dissertation.

Hence, Θ_M of each local model is estimated to minimize the error between $\hat{y}_M$ and y. With this purpose, RLS algorithm detailed in the next section is used.

4.2 Local Model Estimation

Recursive Least Squares RLS algorithm is one the most popular techniques for model parameter estimation and has been continuously optimized through years. In this section it is depicted how this method estimates the vector of parameters for a single operating context, followed by a simulation example in order to best demonstrate its use.

4.2.1 Recursive Least Squares Algorithm

RLS algorithms derivation can be found in many books and papers [58]. For the sake of easy reference, consider a typical RLS algorithm which estimates the best value for the vector of parameters Θ_M by minimizing at each time instant t_k the quadratic error between the model's estimation and plant's consumption, which in the case of the context-based proposed approach is the difference between y and $\hat{y}_M$:

$$J(t_k) = \sum_{t_k \in R} \left(y(t_k) - \hat{y}_M(t_k)\right)^2 \tag{4.8}$$

Where J is the cost function to be minimized, within the boundaries of a region R.

Depending on the order of the regression vector φ from equation 4.2, the Θ_M is comprised of $n_y + n_u$ parameters, as follows:

$$\Theta_M = \left[a_1 \; ; \; \cdots \; ; \; a_{n_y} \; ; \; b_1 \; ; \; \cdots \; ; \; b_{n_u}\right] \tag{4.9}$$

At each time instance Θ_M is adjusted according to its $t_k - 1$ value and a matrix of covariance P_M:

$$\Theta_M(t_k) = \Theta_M(t_k - 1) + P_M(t_k)\varphi(t_k)\epsilon(t_k) \tag{4.10}$$

And P_M is calculated as follows:

$$P_M(t_k) = \frac{P_M(t_k - 1)}{\lambda}\left[I_n - \frac{\varphi(t_k)\varphi(t_k)^T P_M(t_k - 1)}{1 + \varphi(t_k)^T P_M(t_k - 1)\varphi(t_k)}\right] \tag{4.11}$$

A typical value for P_M initialization is $P_0 = I \cdot 10^3$, as in [59], where I is a $n \times n$ identity matrix.

The coefficient λ is denominated forgetting factor. This coefficient may be used to give more importance to recent data ($\lambda \neq 1$) or to take all into consideration ($\lambda = 1$). Values for λ are typically within the range of $0.95 \leq \lambda \leq 1$.

Another similar representation for the forgetting factor, more commonly used in practice is the forgetting window N_λ, which is the number of the most recent samples that are

used for Θ_M estimation:

$$N_\lambda = \frac{1}{1-\lambda} \tag{4.12}$$

For evaluating model's estimation, the average absolute $\overline{|\Delta\epsilon|}$ value of the estimation error ϵ:

$$\epsilon(t_k) = y(t_k) - \hat{y}_M(t_k) \tag{4.13}$$

is most often used. Hence, $\overline{|\Delta\epsilon|}$ is defined as:

$$\overline{|\Delta\epsilon|} = \frac{\sum_{t_k \in R} \epsilon(t_k)}{N} \tag{4.14}$$

Where N is the length of R.

For a better understanding of the RLS method, next section depicts a simulation example, where RLS is used for estimation of a local model on a single operating context.

4.2.2 Simulation Example

In this simulation, a local model is estimated when the plant is operating under a certain operating context. The production variable, plant's consumption and the evolution of RLS parameter estimation are illustrated on figure 4.2.

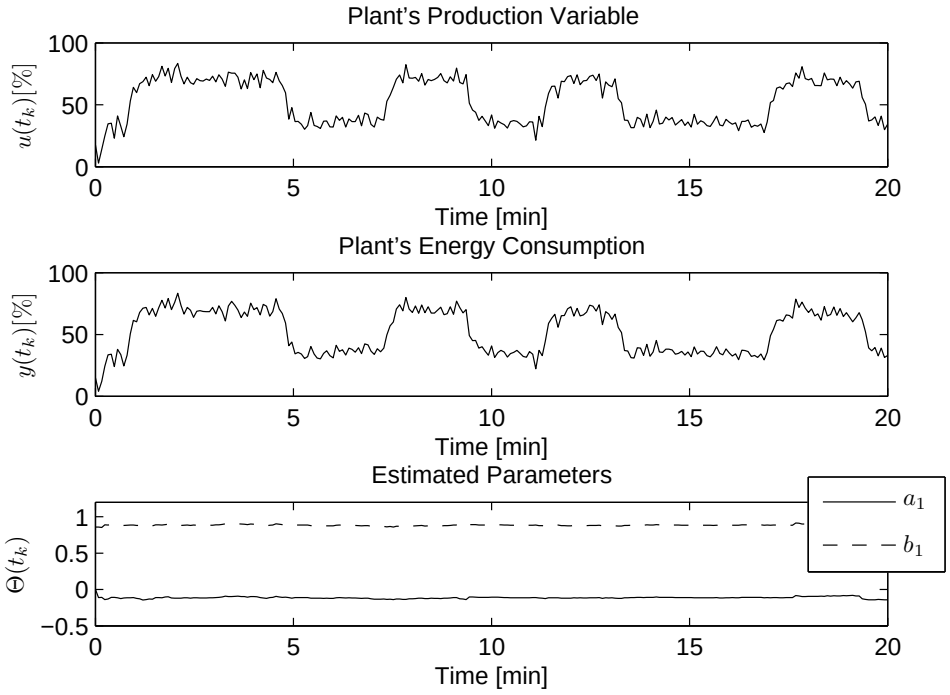

Figure 4.2: Plant's production variable, consumption and estimated parameters.

The RLS algorithm was initialized with a forgetting factor $\lambda = 0.98$, which equals to a sliding window of $N_\lambda = 8$ minutes and the covariance matrix was initialized with $P_0 = I \cdot 10^3$.

The regression vector was chosen with $n_y = 1$ and $n_u = 1$:

$$\varphi(t_k) = [y(t_k - 1)\ ;\ u(t_k)]^T$$

Thus, Θ_M comprises 2 parameters:

$$\Theta_M = [a_1\ ;\ b_1]$$

Thus the local model consumption estimation is:

$$\hat{y}_M(t_k) = a_1 \cdot \hat{y}_M(t_k - 1) + b_1 \cdot u(t_k) + \epsilon(t_k)$$

Figure 4.2 depicts what was already expected. When a single local model is estimated for an operating context, as the plant's behaviour is approximately constant, the estimated parameters rapidly converge to a nearly constant value with low variance, resulting on a consumption estimation with a very low estimation error (in this simulation it is lower than 1.5%), as depicted in figure 4.3

Numerical results of the parameter estimation and the average absolute estimation

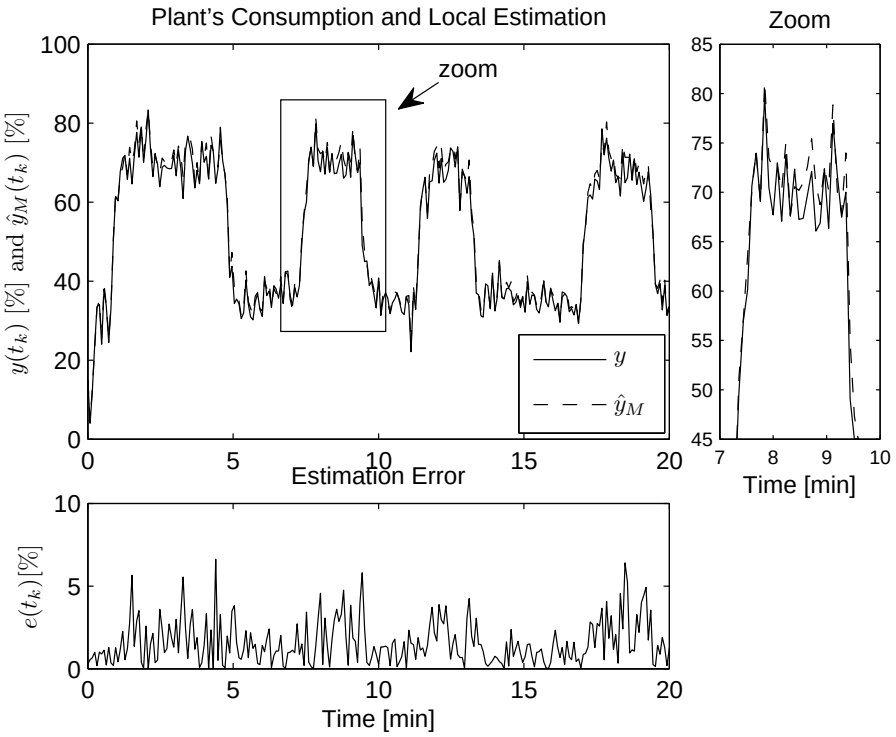

Figure 4.3: Plant's consumption, local model estimation and estimation error

error are shown on table 4.1.

Table 4.1: Estimated Vector of Parameters and Average Absolute Estimation Error

Estimated Parameters
$\Theta_M = [-0.1613\ ;\ 0.8364]$
Average Absolute Estimation Error
$\overline{\lvert\Delta\epsilon\rvert} = 1.46\%$

4.3 Context-Based Multi-model Estimation

As previously discussed, the context-based multi-model approach consists of several models locally estimated according to the identified context. Hence in order to handle multiple models and to able to select when each one should be active for estimation, a structure is proposed in this section.

4.3.1 Bank of Models

As the multi-model approach comprises several local models, it is necessary to distinguish between each one. When a model is selected to be active for estimation, the RLS algorithm estimates the parameters while all the other models are freezed. Therefore, there will be as many RLS algorithms as local models, running in parallel although only one is active at a given time. When a new context is identified, a new RLS algorithm is initialized with the default values. Yet, if the context changes to an operating context in which a model has already been estimated, the RLS algorithm will continue running taking from starting point the estimated parameters and covariance matrix of the already existing model.

Hence, a local model has to incorporate on its structure information regarding the context in which it is estimated. Therefore, a local model is not only identified by its vector of parameters and covariance matrix, but also by the centroid of the operating context. Thus, a local model is comprised of:

- **Estimated vector of parameters** Θ_M — Holds the estimation performed by the RLS algorithm;
- **Covariance matrix** P_M — Holds the covariance matrix for the respective local model;
- **Centroid** c_M — Identifies the operating context in which the model is estimated.

This way, the mapping between context (cause) and consumption (effect) is identified, resolving one of the current problems that state of the art models have.

As models are dynamically created and selected to be active, they have to be stored in a bank indexed by the context's centroid, for further use. The way of how models are selected is depicted in the next section.

4.3.2 Model Selection

When a context is identified by the RT algorithm, a search is performed within the bank to seek for a valid local model in that context. This drives to the question of what determines if a model is valid or not for a certain context.

As depicted previously, RT context identification may result in centroids that are very closer to each other. It was discussed on section 3.2.2 that a suitable interpretation is to let a selection function decide whether or not, those centroids represent the same operating context. Hence, from a multi-model point of view, this may be interpreted as model's sensitivity to changes in context. If sensitivity is "high", small deviations in context centroids are interpreted as a new operating context, thus a new model is created/selected for estimation. On the other hand, if sensitivity is "low" the same model may be selected to be active through a range of centroids, thus stating that those centroids belong to the same operating context.

Therefore, to determine model's validity, it is considered that if the absolute difference between the current context centroid c and an already estimated model's centroid, is not exceeded by a threshold γ (lets name it bank sensitivity) a model is considered to be valid:

$$|c - c_M| < \gamma \tag{4.15}$$

Therefore, upper and lower limits are set for defining model validity, as depicted on Figure 4.4.

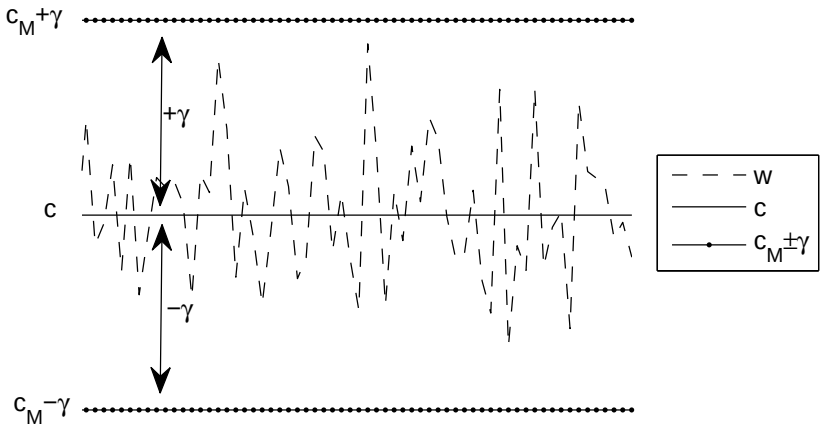

Figure 4.4: Model's upper and lower validity limits near a centroid.

4.3.3 Model Commutation

In order to perform smooth transitions, rather than abrupt jumps between local models estimation when switching from one operating context to another, it is used a transition function as in [60, 61].

The logistic transition function has been commonly used to for this purpose [62], and is defined by the following equation:

$$g(t_k, \beta, p) = \frac{1}{1 + e^{-\beta(t_k - p)}} \tag{4.16}$$

This transition function is bounded between 0 and 100%. The parameter β determines the speed and smoothness of the transition around the changing point p [62], which is defined by the boundary between two operating contexts.

When reaching the changing point, the percentage of model's estimation valid in that region, smoothly decreases from 100%, passing at 50% exactly in the changing point, until it reaches 0%. Whereas, the model valid on the next region has the opposite behaviour, increasing from 0 to 100%, crossing with the first model's estimation on the changing

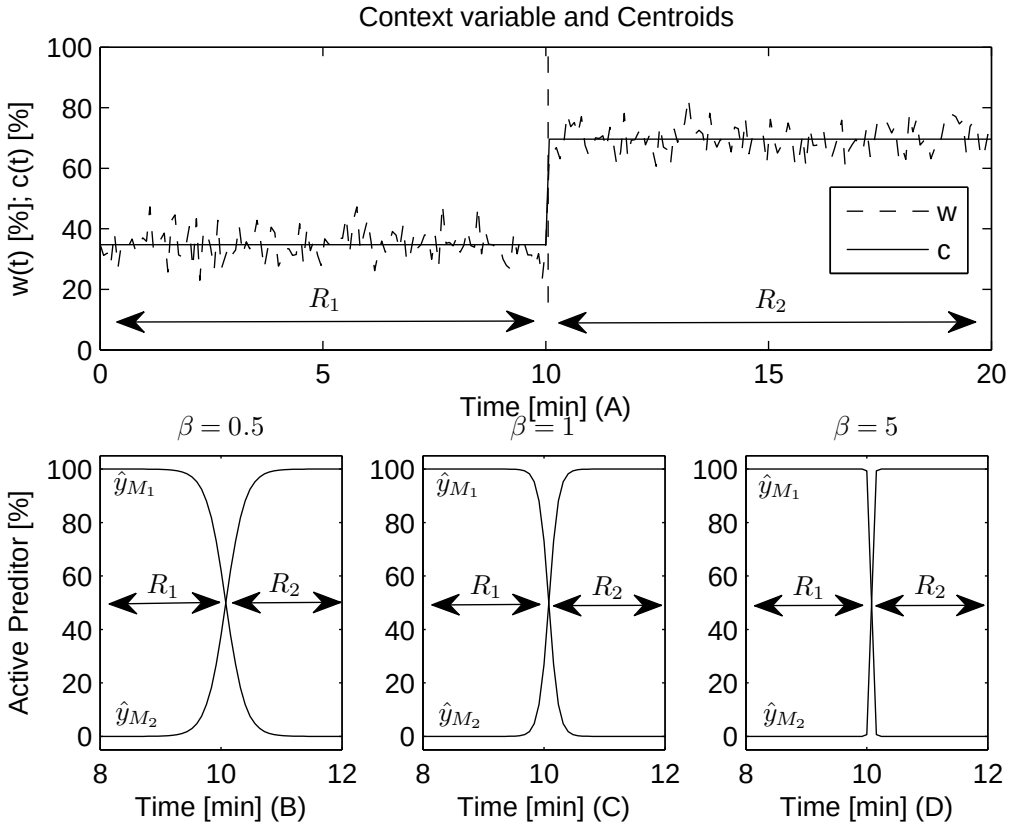

Figure 4.5: (A) Context variable and centroids. Model commutation with (B) $\beta = 0.5$ (C) $\beta = 1$ (D) $\beta = 5$.

point, moment when both models contribute with 50% of their estimation, until it reaches 100%, moment when the commutation is complete.

Hence, the smooth transition reduces the estimation error when changing between context regions, by weighting the contribution of each local model in the estimation near the commutation boundary.

Figure 4.5 (A) depicts an example of a context variable with 2 regions, which consumption is estimated by two local models. In (B), (C) and (D) is depicted the transition smoothness for β values of 0.5, 1 and 5 respectively. As it is depicted, the greater β is, steeper the commutation.

In order to best understand the overall context-based multi-model estimation of energy consumption, consider the following simulation example, which comprises a full demonstration of the concepts depicted until now.

4.3.4 Simulation Example

In this simulation example consider that a plant is operating under the influence of a context variable with a sinusoidal behaviour, ranging between 80% and 30%, commuting approximately every 11 minutes. Production, energy consumption and context variable data are depicted in figure 4.6, respectively in (A), (B) and (C).

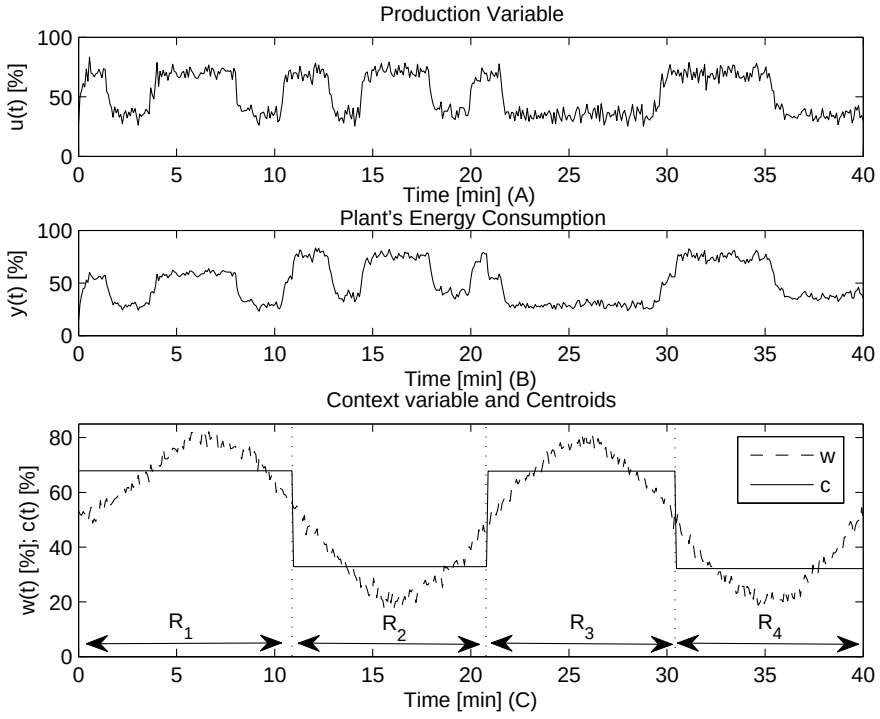

Figure 4.6: (A) Plant's production data. (B) Energy consumption data. (C) Context variable, regions and centroids.

The plant's energy consumption is simulated by the following expression:

$$y(t_k) = a_1 \cdot y(t_k - 1) + b_1(w(t_k)) \cdot u(t_k) + \epsilon(t_k)$$

For ease of demonstration, consider that only one parameter is context dependent and is given as follows:

$$b_1(w(t_k)) = \begin{cases} 1 + \frac{w(t_k)}{\iota} \cdot b_1 & , w(t_k) \leq 50\% \\ \frac{3}{2} + \frac{w(t_k)}{\iota} \cdot b_1 & , w(t_k) > 50\% \end{cases}$$

Where $b_1 = 0.6321$ is the constant parameter in the absence of context influence. The parameter a_1 is a coefficient given by $a_1 = -0.3679$. The constant ι[2] is an impact adjustment coefficient, changing this value will change the magnitude of the impact that the context variable has on the parameter b_1.

The RLS algorithm is initialized with a forgetting factor $\lambda = 0.98$, covariance matrix $P_0 = I \cdot 10^3$ and to match plant simulation, the estimated parameters are $\Theta = [a_1 \; ; \; b_1]$.

Regarding, the RT algorithm, it is initialized with $\tau = 10$ minutes to approximately match the context variable commutation period and with $\kappa = 10^{-6}$, resulting in the partition of the context variable into 4 regions as depicted in figure 4.6 (C) and shown on table 4.2.

Although the regression tree analysis identified 4 regions, region's centroids c_{R_1} and c_{R_3} are very similar, as both correspond to the upper half of the context variable, differing in less than 0.1%. Same happens regarding regions R_2 and R_4.

Therefore, in order to consider that R_1, R_3 belong to the same operating context, a context tolerance γ of 1% is enough for the bank of models to recognize that the model to be estimated in the context region R_3 is the same as the one in R_1, and similarly the model to be estimated in R_4 is the same as R_2.

Figure 4.7 (A) depicts which model is active for estimation and the validity range for each one with a $\gamma = 1\%$. As it is depicted, model M_1 is estimated on the first context region R_1. When the context region changes to R_2 a new RLS algorithm for model M_2 is initialized as it is perceptible by the increased variance on the estimated parameters as depicted in (B).

Table 4.2: Regions, Centroids and Region Boundaries.

Region	**Centroid [%]**	**Region Boundaries [min]**
R_1	$c_{R_1} = 67.90$	$0 \leq t < 10.55$
R_2	$c_{R_2} = 32.86$	$10.55 \leq t < 20.50$
R_3	$c_{R_3} = 67.74$	$20.50 \leq t < 30.26$
R_4	$c_{R_4} = 32.17$	$30.26 \leq t < 40$

[2]Note that ι is only used for simulation purpose.

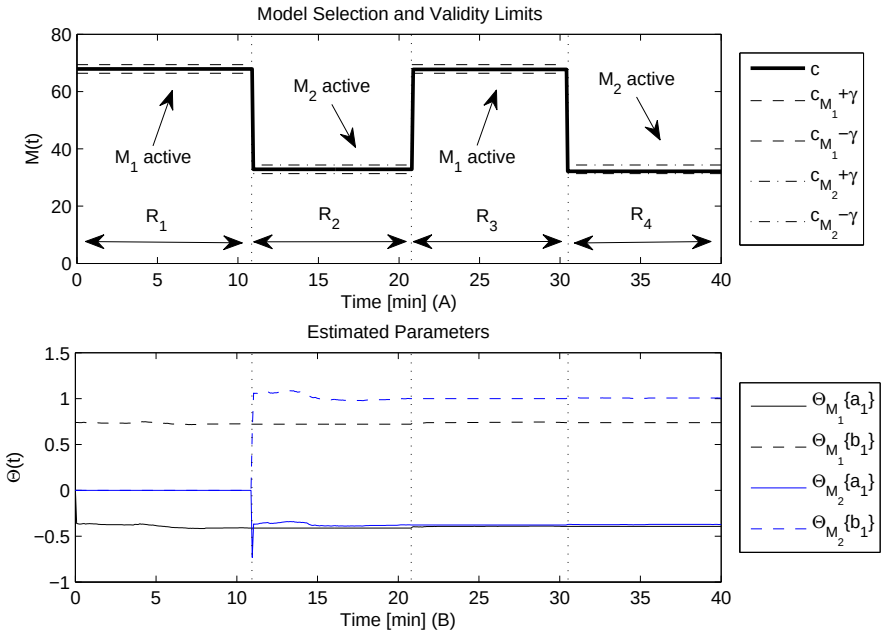

Figure 4.7: (A) Active model and validity limits. (B) Estimated parameters.

When the context region changes to R_3 instead of initializing a new RLS algorithm, the bank of models recognizes that the respective region's centroid is within the validity range of model M_1, hence the algorithm continues running from the parameters and covariance matrix of M_1. This effect is perceptible by the slow variance of the parameters being estimated within R_3. Same happens with region R_4, instead of estimating a new model, the RLS algorithm continues the estimation process from the model M_2.

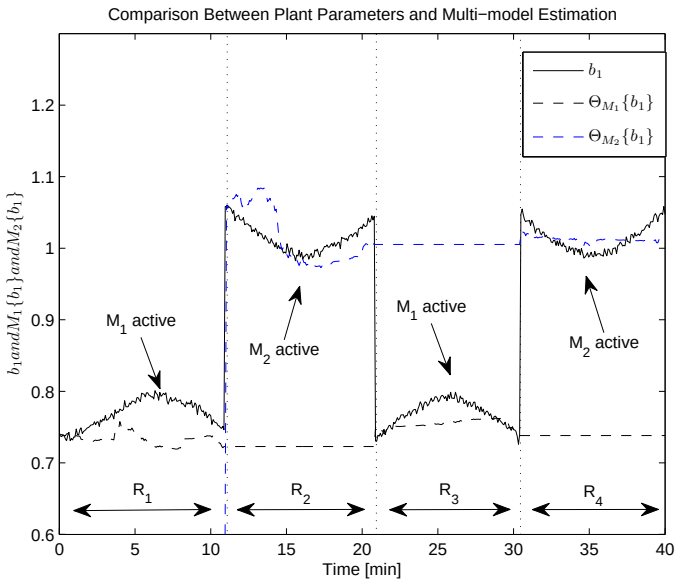

Figure 4.8: Parameter estimation for the upper half and lower half of the context variable.

As it is expected both models estimate the a_1 very closely, as this parameter is not affected by the context variable. However, b_1 shows a significant difference between the two models as this parameter is affected by the context variable. Figure 4.8 depicts a comparison between the plant's parameter b_1 and the estimation by both local models. As it is shown, $M_1\{b_1\}$ estimates b_1 when the context variable is within its upper half, whereas $M_2\{b_1\}$ estimates in the lower half. This partition, provides a much better estimation than a single model trying to estimate b_1 through all plant's operation (traditional approach).

Thus, the result is a bank which contains two local models: M_1 and M_2:

$$\hat{y}_{M_1}(t_k) = \Theta_{M_1}\{a_1\} \cdot \hat{y}_{M_1}(t_k - 1) + \Theta_{M_1}\{b_1\} \cdot u(t_k) + \epsilon(t_k)$$

$$\hat{y}_{M_2}(t_k) = \Theta_{M_2}\{a_1\} \cdot \hat{y}_{M_2}(t_k - 1) + \Theta_{M_2}\{b_1\} \cdot u(t_k) + \epsilon(t_k)$$

Therefore, the global context-based multi-model is:

$$\hat{y}(t_k) = \begin{cases} \hat{y}_{M_1}(t_k) & , |c - c_{M_1}| < 1\% \\ \hat{y}_{M_2}(t_k) & , |c - c_{M_2}| < 1\% \end{cases}$$

Figure 4.9 (A) depicts the multi-model estimation and plant consumption. In (B) it is depicted the commutation between local models using the logistic transition function with $\beta = 1$.

As it is depicted, although the context variable influence, the estimation is very closer

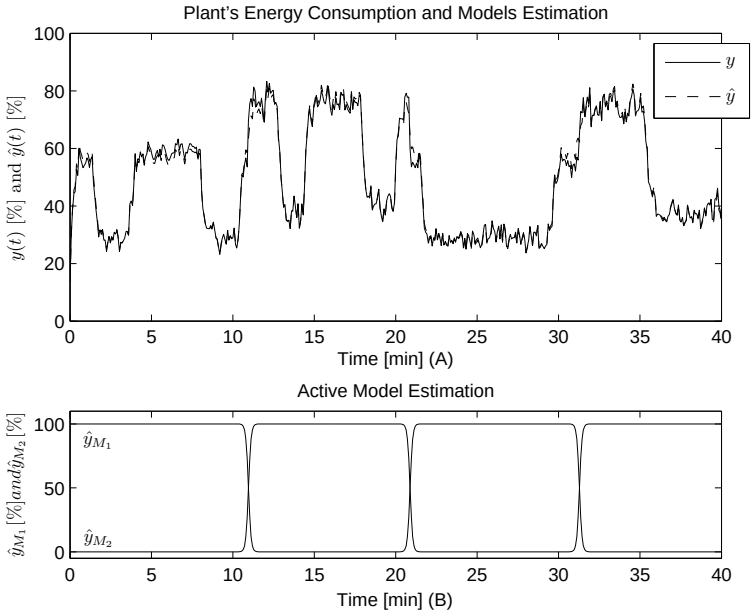

Figure 4.9: (A) Plant energy consumption and CBEM estimation. (B) Local models commutation.

Table 4.3: Models, respective centroids and estimated parameters.

Model	Centroid [%]	Parameters
M_1	$c_{M_1} = 67.90$	$\Theta_{M_1} = [-0.3940\ ;\ 0.7374]$
M_2	$c_{M_2} = 32.86$	$\Theta_{M_2} = [-0.3732\ ;\ 1.0047]$

to plant's consumption through the entire plant operation. This is achieved by the commutation of local models and automatic identification of the operating context. The average absolute estimation error of the multi-model estimation is of $\overline{|\Delta\epsilon|} = 1.08\%$, which shows an accurate consumption estimation.

Numerical results of the estimated parameters for models M_1 and M_2, along with respective model's centroids are shown on table 4.3

For establishing a comparison for the proposed approached performance, consider in the next section the same simulation performed with the traditional approach.

4.3.5 Comparison with Traditional Approach

This simulation has the purpose of showing the benefits of incorporating context information in energy consumption estimation, by means of a comparison with the traditional approach applied to the same scenario as the previous example 4.3.4, in which a plant operates under the influence of a context variable. Results for both approaches are compared.

As described previously the traditional approach consists of a single model (hence a unique RLS algorithm) to estimate consumption through all plant's operation, disregarding context information. Thus, this approach only focus on identifying the relation between production and consumption data, not establishing any connection with the context in which the plant is operating. Lets consider the nomenclature M_t for referring to the model estimated by the traditional method and in contrast M_{cb} for referring to the context-based approach.

For comparison purposes, consider that the RLS algorithm is initialized with the same configuration as on the previous example, for both approaches.

Figure 4.10 depicts the comparison between the plant's parameters, the context-based multi-model M_c estimated parameters and the traditional model M_t estimation. Although the plant's parameter b_1 is strongly influenced by the operating context, $\Theta_{M_{cb}}\{b_1\}$ estimation is able to follow the influence. On the other hand, the parameter $\Theta_{M_t}\{b_1\}$ estimated by the traditional approach, strongly fails when the context changes, as it is perceptible starting on 11^{th} minute.

In an attempt to compensate the influence on parameter b_1, not only $\Theta_{M_t}\{b_1\}$ is affected, but also $\Theta_{M_t}\{a_1\}$, leading to a failed estimation on both parameters. As the traditional approach uses a single RLS algorithm, the influence of the operating context produces poor results in the estimation, since the tracking is lost due to outdated samples from a previous context region, which contribute to the estimation when the context

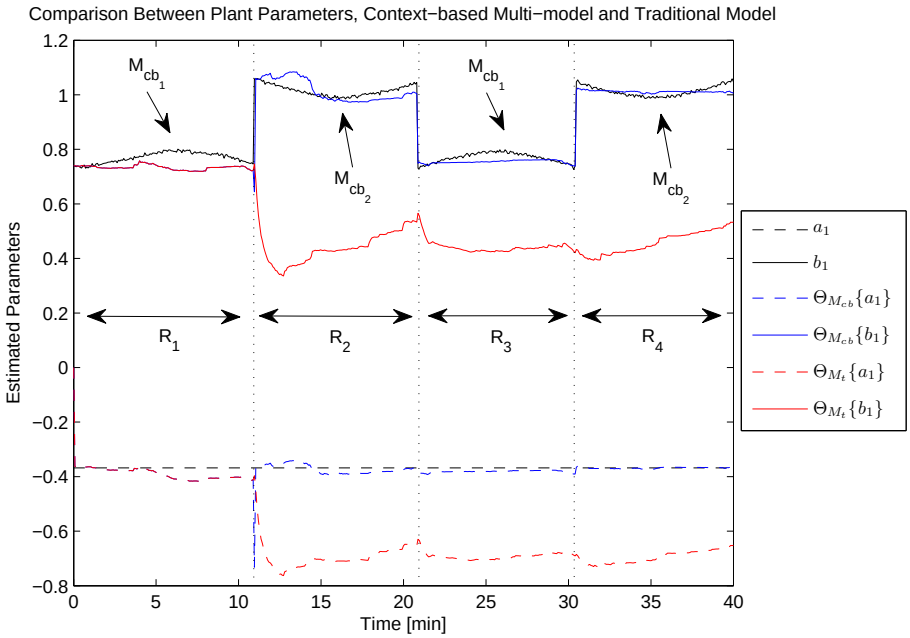

Figure 4.10: Comparison between CBEM parameter estimation, tradition approach estimation and plant parameter.

changes. Although adding a forgetting factor may reduce the importance of the outdated samples, reducing significantly the forgetting window may turn the algorithm unstable, as in [63]. Another important drawback, is that the traditional method is based on a single model for the entire plant's operation, thus loss of information about how context influence energy consumption is lost as there is no connection between the operating context and the model estimated.

The context-based approach is able to keep track of the context influence, given that a new model is created when a new operating context is identified. When a new model is created, the covariance matrix is initialized, allowing a rapid convergence of the parameters [63] as the outdated samples are discarded, not impacting the estimation within the new context region. Also, as a new model is created for each operating context and identified by the respective centroid, it is possible to use track down causes to consumption patterns

Figure 4.11 (A) depicts the comparison between the plant's energy consumption, the context-based and traditional approach predictions. On (B) is depicted the estimation errors for both approaches.

As it is depicted, the context-based multi-model approach is able to follow the plant's energy consumption closely through the entire plant operation. Whereas, the traditional approach fails to estimate consumption, specially between the time-ranges from the beginning of the simulation until the 11^{th} and from the 21^{th} to the 31^{th} minute, which are the boundaries of context regions R_1 and R_3, respectively.

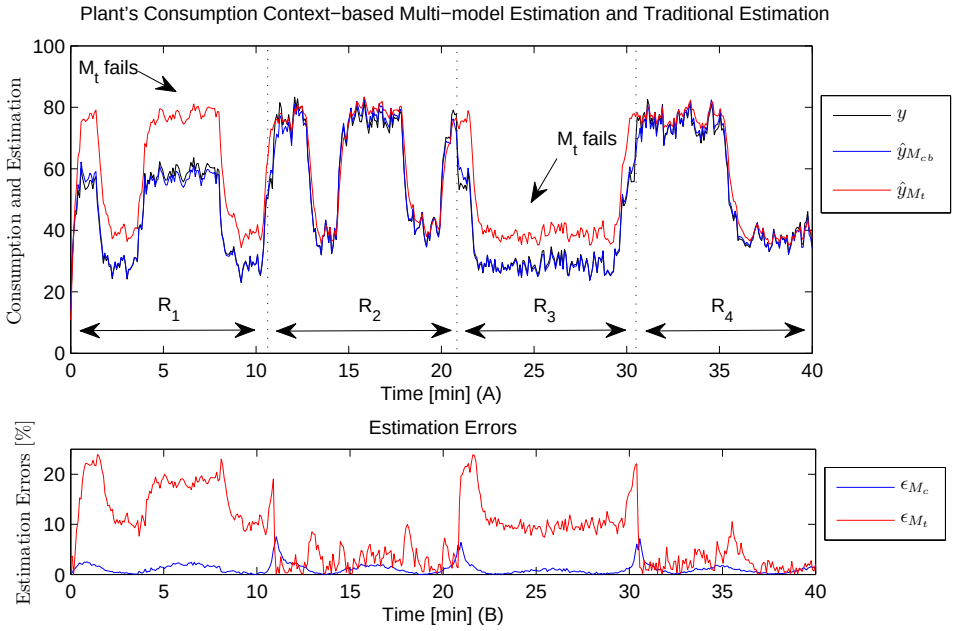

Figure 4.11: (A) Comparison between plant energy consumption, CBEM and traditional approach estimation. (B) Estimation error for both approaches.

This is due to the use of a single RLS algorithm. As the last samples taken into consideration for estimation correspond to the samples within R_4, the traditional model is closer to estimate consumption on regions R_2 and R_4, in which consumption is affected in the same way, than on the others. Thus, the consumption pattern observed on regions R_1 and R_3 is not explained by the traditional method.

The average absolute estimation error of the context-based approach is considerably lower (approximately 8 times less) than the estimation error from the traditional approach. Numerical results for average absolute estimation error are shown on table 4.4.

This simulation, although theoretical, shows the benefits of the context-based approach, which particularly improves energy consumption estimation (compared to the traditional approach) and more importantly it depicts a method for relating context information (providing the causes) with consumption patterns.

Table 4.4: Comparison between traditional and context-based average absolute estimation error.

Context-based Multi-Model Approach Estimation Error
$\overline{\lvert\Delta\epsilon\rvert} = 1.08\%$
Traditional Model Approach Estimation Error
$\overline{\lvert\Delta\epsilon\rvert} = 8.27\%$

5

Implementation

In this chapter it is proposed an architecture to implement the context-based energy consumption identification system. The architecture is described, detailing each individual block function, inputs and outputs.

A simulator is also presented in order to gain more sensitivity with the context-based approach and to allow simulations and experimentations with different system initializations.

5.1 Developed System

5.1.1 Methodology

Context-based identification of energy consumption involves four different stages.

The first stage is where measurements are made in order to retrieve information about plant operation. Production variables, energy consumption and context data are sampled. In this stage it is important to establish essential relations between physical processes that play a key role on plant's operation and energy consumption patterns that they present while performing specific tasks. It is also of extremely importance to correctly select which context variable should be taken into consideration of context identification.

The second stage is where context variables are partitioned into context regions. Each context region may represent an operating context, which is identified by a unique centroid, or the union of centroids in case of multiple context variables. Centroids are the link between local models and operating contexts. The partition of context variables into regions was discussed on chapter 3.

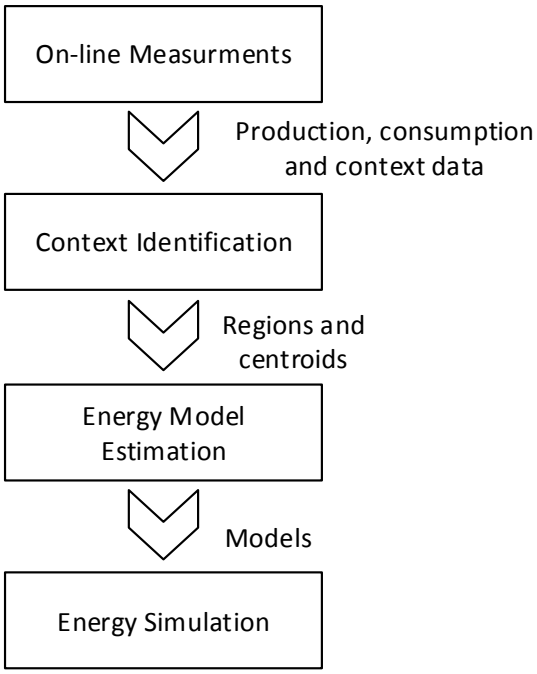

Figure 5.1: Schematic illustration of context-based identification stages.

The third stage comprises the estimation of the context-based consumption models, which comprise several local models, each one for an operating context. According to the current context region, a model is selected from the bank of models and used to estimate the local energy consumption dynamic. The method for selecting the appropriate local model and the algorithm to estimate model parameters was previously described on chapter 4.

The last stage is where simulations are made using the estimated models. As models are directly linked to the operating context, further analysis may be performed over the models, in order to better understand how context variables influence energy consumption or to predict energy use under certain pre-defined context conditions. A schematic illustration of the overall approach is shown in figure 5.1.

5.1.2 Architecture

Following the discussion on the previous section, consider the proposed architecture depicted on figure 5.2, which implements the methodology presented.

The block identified as Plant represents the plant consumption in whole or a single production unit. While the plant is operating, it is exposed to context variables which influence its behaviour. Process data and energy consumption patterns are continuously measured and stored on a regressor vector to be used for model estimation. Context variables are also continuously measured. Blocks identified as Production data and Context

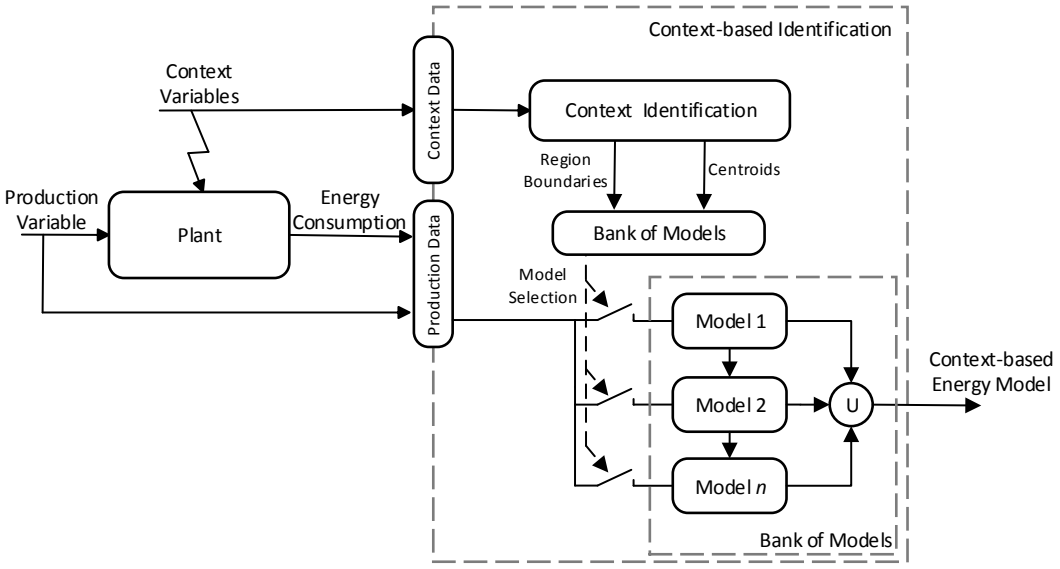

Figure 5.2: Proposed architecture for the context-based energy consumption identification system.

data hold, respectively, the production and context samples.

The block identified as Context Identification is responsible for the partitioning of context variables into context regions and calculate their respective centroids and boundaries. This is performed by applying regression tree analysis over the context variables samples. The output of the Context Identification block is a set of rules, that can be represented as a tree diagram, which define region boundaries and respective centroids.

The block denominated as Bank of Models is responsible for selecting a model to learn local energy consumption and to hold information about which model is related to an operating context by its centroid. The model is selected according to the proximity of its centroid and the current context region centroid. Only one model is allowed to estimate energy consumption at a time, using a RLS algorithm for that purpose, while all the others are freezed. Hence, the bank of models is comprised of multiple local models each one linked to the respective operating context centroid.

Next section describes in detail the functional analysis in terms of sequential steps performed by the context-based energy consumption system.

5.1.3 Functional Steps

Figure 5.3 depicts the functional diagram of the context-based energy consumption identification system.

While production and context data are being measured from plant operation, the context variables are decomposed into a set of context regions and for each one, centroids and

regions boundaries are calculated.

When a new context region is detected, a search is performed in the bank of models in order to select the model that is valid to learn the local energy consumption for the current region. A model is considered to be valid if its centroid validity limits contains the centroid of the current region. If a valid model is found, parameters and covariance matrix are retrieved to continue estimation. If not, a new RLS algorithm is initialized, thus a new model is selected to be active and to start estimation. The same data regressor vector is shared between RLS algorithms, although only one model is active at a time, while all others are freezed.

This loop is performed until no more context regions exist. At this point it is requested more data, and the process starts again.

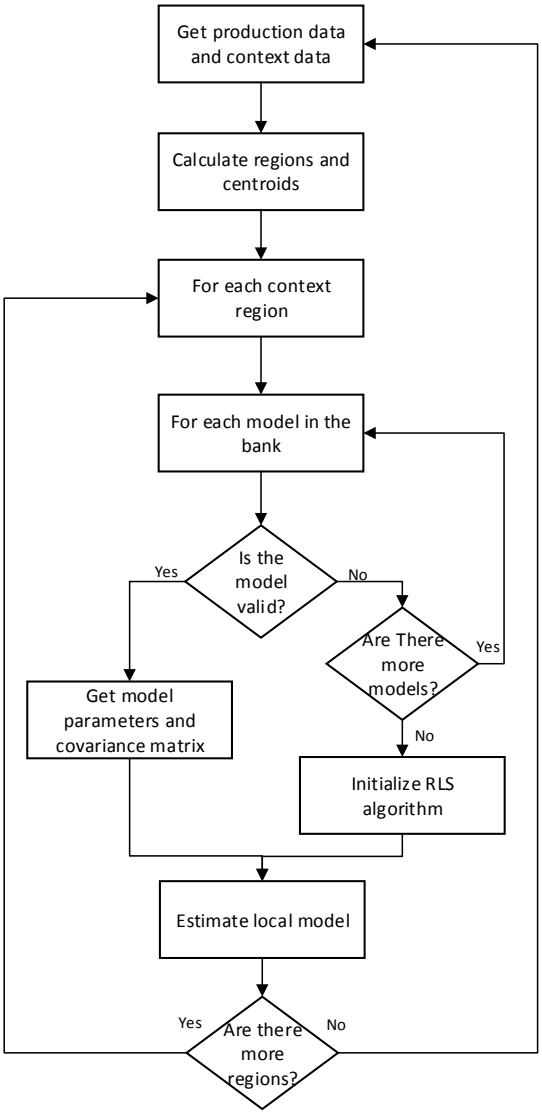

Figure 5.3: Functional diagram of the context-based energy consumption system.

5.2 Context-Based Identification Simulator

5.2.1 Global Overview

In order to demonstrate the context-based energy consumption approach, the system described in the previous chapter is implemented in a computational simulator developed in Matlab®. The simulator main goal is to create a user friendly interface able to run the developed system. The user is able to change algorithms specifications, initializations and analyse the results at the same time, providing the ability to better tune the algorithms. The computational platform Matlab® was chosen to develop the interface, given the wide use between researchers and universities, due to the optimized methods available and easy to use tools, specially in signal processing and dynamic systems areas. Figure 5.4 depicts the front-end interface.

The simulator interface comprises 5 blocks and is able to run either in simulation or real experiment mode. A brief description of each blocks function is described as follows:

- **Block A - Simulation and experiment specifications** — This block comprises simulation and experiment specifications, such as generating production and context data in case of simulation mode or import external data in case of experiment mode, allowing the user to run real experiments;

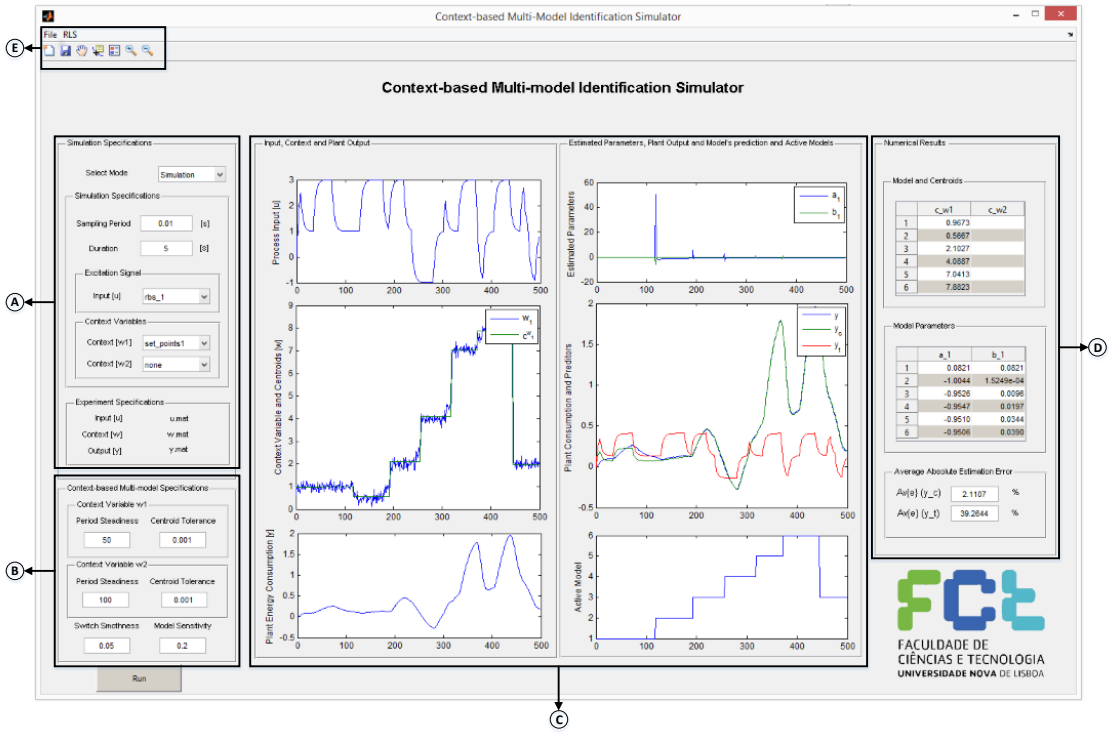

Figure 5.4: Context-based energy consumption simulator interface.

- **Block B - CBEM specifications** — This block comprises all related CBEM algorithms specifications and initializations, both for context identification (RT algorithm) and bank of models (RLS algorithm) specifications.
- **Block C - Graphical results** — This block displays graphical information regarding production, context data and CBEM results, such as centroid calculation, estimated model parameters, active models and model predictions.
- **Block D - Numerical results** — This block displays the numerical results of the developed system. Numerical values of model's parameters, centroids and prediction errors are displayed.
- **Block E - Menu and toolbar** — This block comprises extra available menus, and tools. The menu RLS is used for configuring the RLS algorithm initialization. Other tools for handling the simulation are also available, such as save, restart and managing figure's information.

The following sections, describe in more detail each block and all the functionalities that are available for user interaction.

5.2.2 Block A - Simulation and Experiment Specifications

Figure 5.5 depicts the simulator, zoomed on block A region for a better visualization.

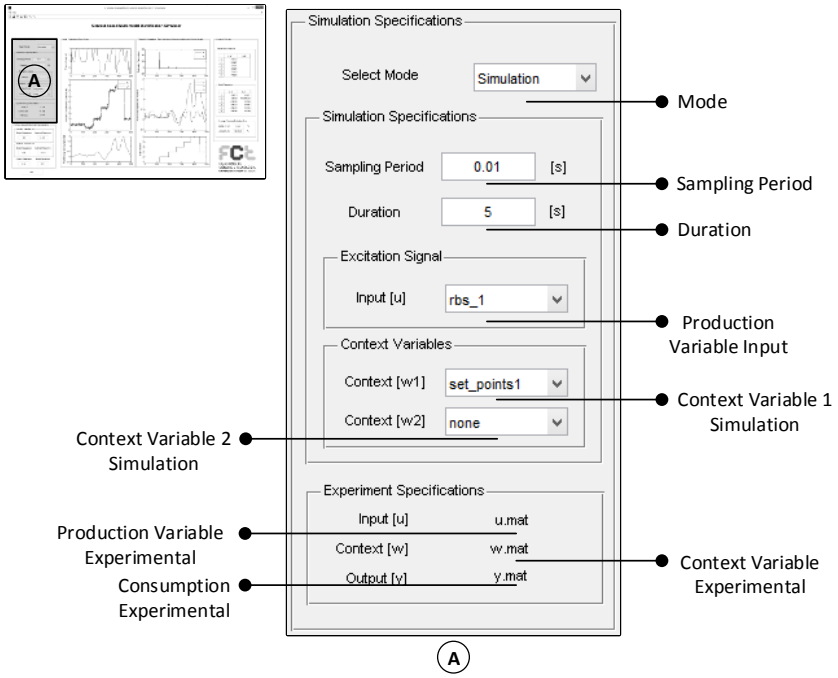

Figure 5.5: Simulator interface zoomed on the block A region.

This block allows the user to select specifications regarding the simulation or experiment to be performed. The simulator is developed to work on a simulation or experiment mode. The drop-down list [1] named Mode allows the user to choose between simulation or experiment.

In the simulation mode, the user is able to define the sampling period and duration, both in seconds, of the simulation by editing the edit boxes[2] denominated as Sampling Period and Duration, respectively. In this mode, the user is also able to select plant's production data , by selecting one of the signals available, using the drop-down list denominated as Input Signal. Similarly, the user is also able to define context variables. For simulation purpose, only 2 context variables are processed simultaneously, and can be chosen by selecting the desirable signal from the drop-down lists Context Variable 1 Simulation and Context Variable 2 Simulation.

Default signals already built-in the application are:

- **Impulse** — An impulse signal, ideal to simulate the impulsive response.
- **Step** — A signal which magnitude commutes from 0 to 1 at the middle of the simulation;
- **Ramp** — A signal that increases its magnitude during the simulation time.
- **Multiple set-points** — A signal that is piece-wise constant over several periods of the simulation.
- **Random binary sequence** — A signal which comprises several random binary sequences, ranging between 0 and 1;
- **Constant** — A signal that is constant over the entire simulation;
- **Sinusoidal** — A sinusoidal signal. Default value of frequency is automatically adapted in order to show a complete period over the simulation;

According to the sampling period and duration values, signals automatically adapt to the simulation specifications. The user may also import a signal by selecting the option import.

Regarding plant simulation, a difference equation is used to simulate energy consumption. The difference equation coefficients are available in a configuration file for the user to change if desired, allowing the user to use and simulate the context-base identification approach with a wide range of dynamic systems.

If the experiment mode is selected, all options described above are locked. In the experiment mode process data, context variables and energy consumption data are directly imported by loading the `.mat` files shown on the simulator interface, respectively: `u.mat`; `w.mat`; `y.mat`. Each signal should also be followed by a time vector.

[1] A drop-down list is a graphical control element, that allows the user to choose one value from a list.

[2] An edit box is a graphical control element, that allows the user to input text to be used by the software.

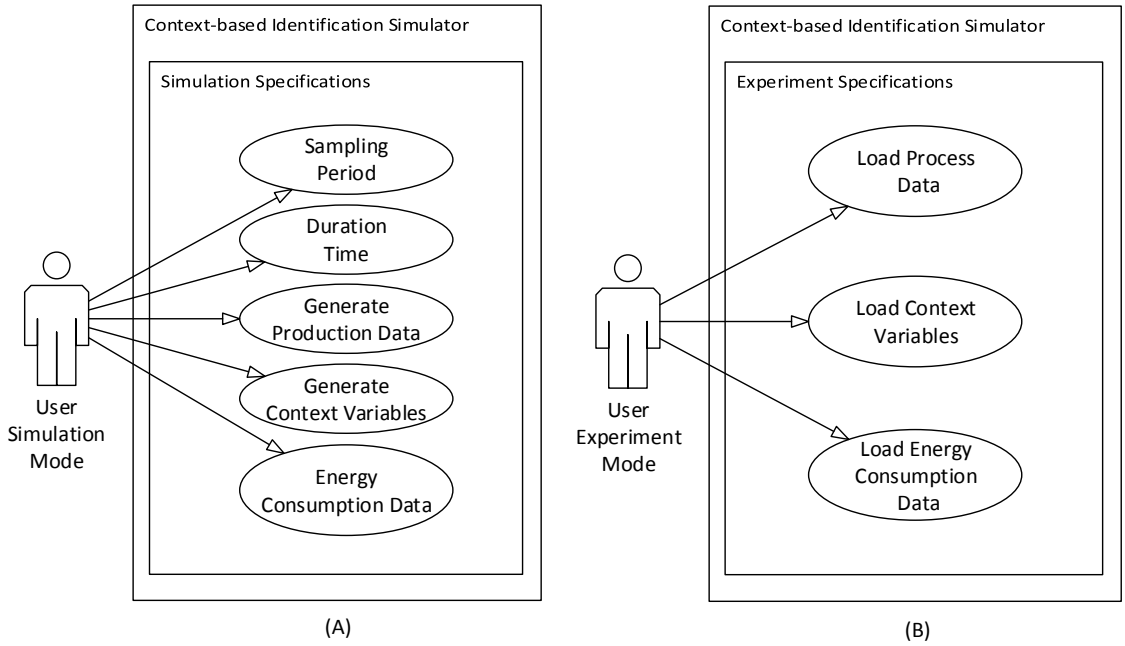

Figure 5.6: Use case UML diagram for (A) simulation and (B) experiment mode.

The use case diagrams for both simulation and experiment mode functionalities are depicted on figure 5.6 (A) and (B) respectively.

5.2.3 Block B - Context-based Multi-Model Identification Specifications

Figure 5.7 depicts the simulator zoomed on the block B region.

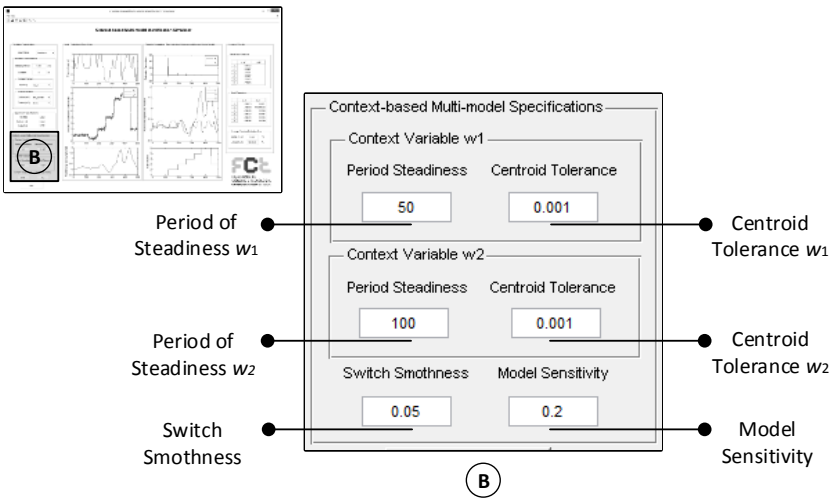

Figure 5.7: Simulator interface zoomed on the block B region.

This block allows the user to configure regression tree algorithms and the bank of models specifications.

For each context variable, a regression tree algorithm is used to partition context variables into regions and calculate centroids. The degree of partition is dependent on the initial values of the algorithm. On section 3.2.1 it was presented the definition of centroid tolerance κ and minimum period of steadiness τ, which are require for algorithm initialization. Using the edit boxes denominated as Period of Steadiness and Centroid Tolerance, the user is able to play with the algorithm initialization and perform simulations to analyse the impact that this initializations have on the context variable partition. Plus, the user is also able to best tune the algorithm for the type of simulation that is performing, increasing the quality of the entire system.

Regarding the initialization of the bank of models, it was presented on section 4.3.2 the definition of model validity limits, which are set according to a initialized threshold denominated as bank sensitivity γ. Changing the bank sensitivity will result on an increased or decreased sensitivity of the bank regarding context regions centroids, as each model validity limit is changed. The user can initialize the sensitivity of the bank by editing the edit box denominated as Model Sensitivity. Also, the transition between local models was detailed on section 4.3.3, defining the concept of transition smoothness, which is dependent on the coefficient β denominated as smooth coefficient. This coefficient is available for the user to initialize, by editing the edit box named Switch Smoothness.

The use case UML diagram regarding user available functionalities is depicted on figure 5.8.

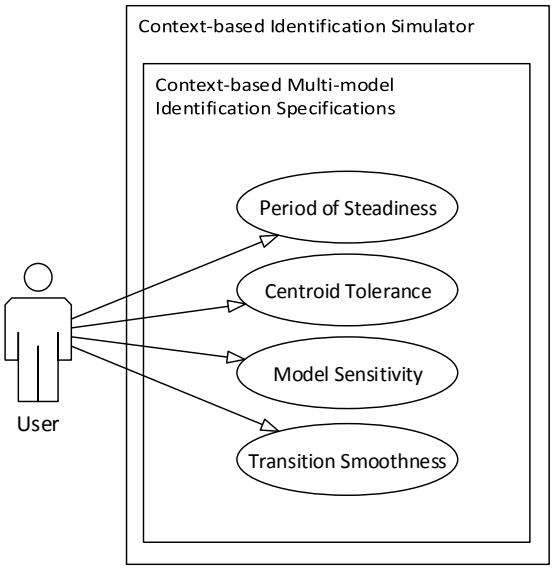

Figure 5.8: Use case UML diagram for context-based mode multi-model identification specifications.

5.2.4 Block C - Graphical Results

Figure 5.9 depicts the simulator, zoomed on the block C region.

In this section, the user is able to visualize the graphical results of the developed system after the algorithm is executed. The time-series plot[3] of production data and plant energy consumption is shown on the plots denominated as Production Variable and Plant Energy Consumption, respectively. These plots are generated according to the user selection of signals as described on the section 5.2.2, where block A was detailed. They can either represent simulations, or experiments according to the mode previously selected and signals chosen.

Concerning context identification, context variable and centroids are also shown on the plot named Context Variables and Centroids. In this plot the context variable signal is shown and overlapped with the regression tree output, allowing the user to compare centroids and region boundaries with the context variable.

Regarding energy model estimation, the user is able to visualize the evolution of parameter estimation over time by analysing the plot denominated as Estimated Parameters. These parameters are the result of each RLS algorithm, in which the vector of parameters is updated. Also, the model that is select to be active is shown on the plot called Active Models.

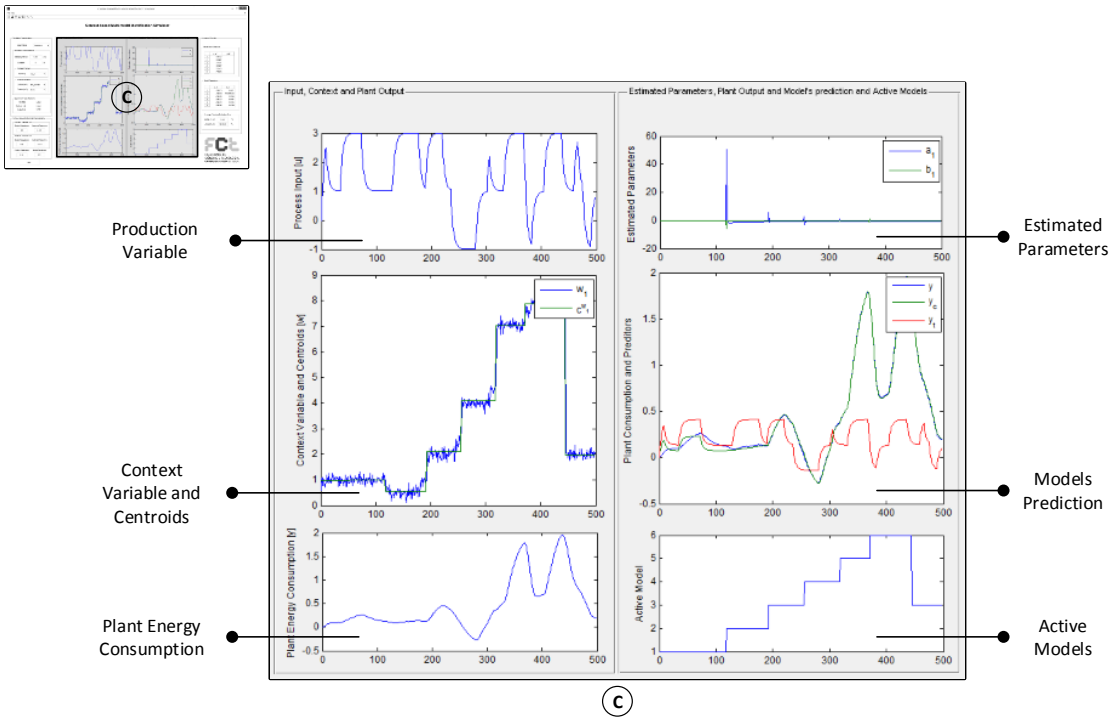

Figure 5.9: Simulator interface zoomed on the block C region.

[3]In graphical interface, a plot is a display element used to create and display graphical representations of data.

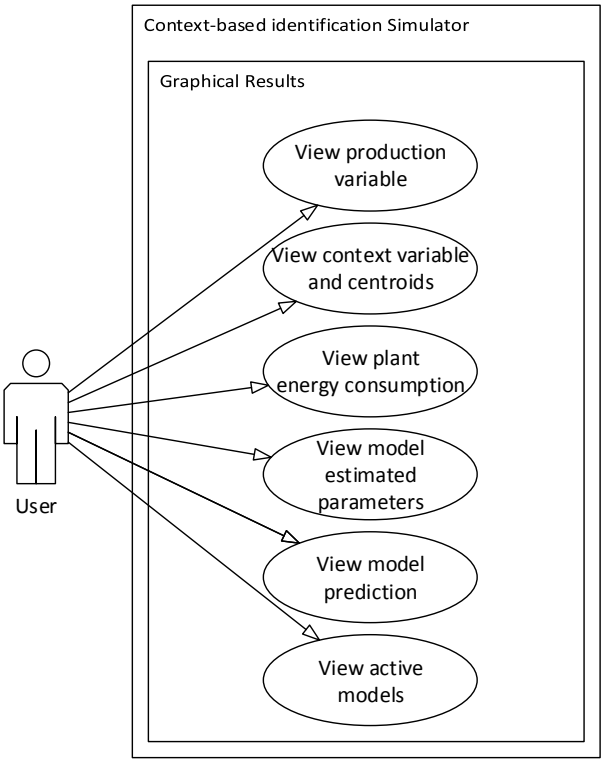

Figure 5.10: Use case UML diagram for graphical results.

These informations, combined with context variables and centroid calculation, provide a better understanding of the context-based identification system performance according to the selected specifications/initializations.

Finally, a comparison between plant's energy consumption, context-based energy model and the traditional method estimation, similar to the comparison depicted on section 4.3.5, is shown on the plot called Models Prediction. The overlap of both approaches with the plant energy consumption allows the user to compare the performance of each one and also highlights the benefits of the proposed context-based approach when compared to the traditional method.

The use case diagrams and user allowed functionalities are depicted on figure 5.10.

5.2.5 Block D - Numerical Results

This section comprises the numerical results of the simulation. Figure 5.11 depicts the simulator, zoomed on block D region for a better visualization.

During the simulation, the local models and respective centroids are displayed on the table[4] denominated as Bank of Models. The bank of models may be loaded from a previous simulation, as it will be discussed on the next section, or it may be the result of

[4]In graphical interfaces, a table is an element control used to display data in a structured form, separated by columns and rows.

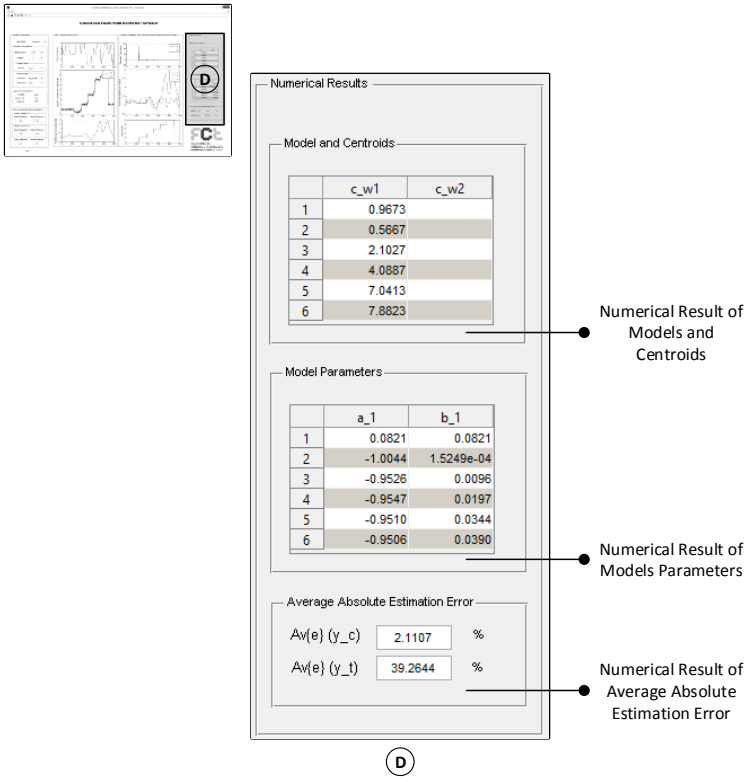

Figure 5.11: Simulator interface zoomed on the block D region.

the current simulation. Either ways, the bank that is currently in use is displayed.

Similarly, each model's estimated parameters are shown on the table called Model Parameters. According to the order of the model, which is defined at the initialization time of the RLS algorithm, the table is adapted in order to display all parameters that comprise the model.

This information is very useful to understand the way of how models are created and to give a global view of the bank of models status and evolution through simulations.

Concerning system performance, the average absolute estimation error of the context-based energy model is displayed on the text-box [5] identified by the label[6] $Av\{e\}(y\_c)$, where y_c is the global context-based model predictor, resultant of the union of all local model predictors. Likewise, the average absolute estimation error of the traditional approach is also displayed, in this case on the text box called $Av\{e\}(y\_t)$.

The numerical comparison between both approaches is a complement of the graphical comparison described in the previous chapter, particularly regarding each method accuracy on the estimation of consumption.

As a summary, figure 5.12 illustrates the use case UML diagram, displaying all features that are available for the user.

[5] A text-box is a graphical element, intended for displaying information, that only allows the user to view text but not edit.

[6] In graphical interfaces, labels are elements used to identify control elements.

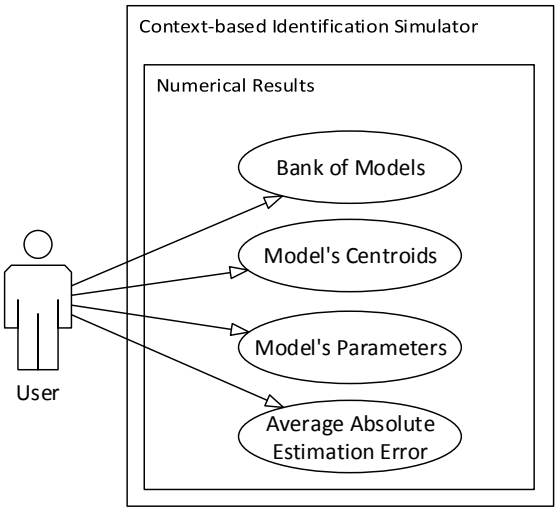

Figure 5.12: Use case UML diagram for numerical results.

5.2.6 Block E - Menu and Toolbar

This section allows the user to use software related tools, which are spread between the Menu bar[7] and Toolbar[8], as depicted on figure 5.13 (A).

The toolbar allows the user to use quick shortcuts to start a new simulation or to save the current one. Also, the toolbar provides the user some tools to better exploit plots:

- **Move Plot** — Allows the user to navigate forward and backwards on the data that is shown on a plot;

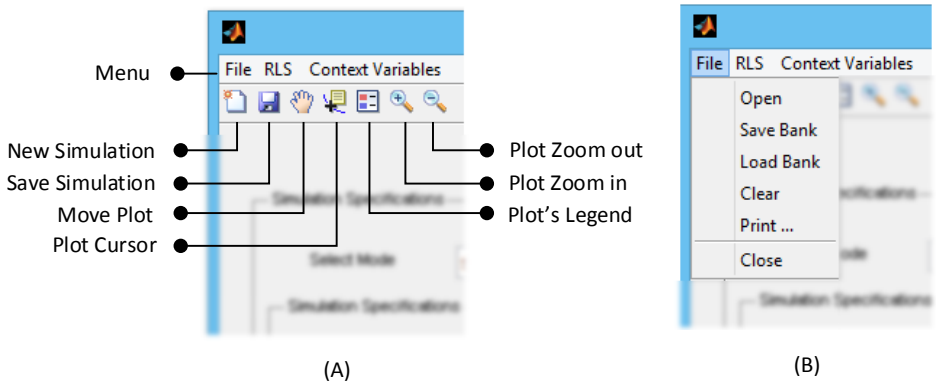

Figure 5.13: (A) Menu bar and toolbar. (B) File menu options.

[7]In graphical interfaces, a menu bar is a control element which contains application-specific menus which provide access to functions or new windows.

[8]In graphical interfaces, a toolbar is a graphical control element on which on-screen buttons are placed.

- **Plot Cursor** — Allows the user to select a specific point on a plot, showing specific data values for that point;
- **Legend** — Disables or enables plot legends;
- **Zoom in and Zoom out** — Allows the user to better navigate on the plot by zooming in or out.

Regarding the menu bar, a file menu, RLS menu and context variables menu is available. The file menu, as depicted in figure 5.13 (B) provides to the user the following abilities:

- **Open** — Opens a simulation previously saved, loading all specifications and data.
- **Save Bank** — Saves the current bank of models.
- **Load Bank** — Loads an existing bank of model.
- **Clear** — Clears all plots.
- **Print** — Prints the simulator current state.
- **Close** — Closes the simulator.

The RLS menu is depicted on figure 5.14. This menu is specific for the RLS algorithm initialization and allows the user to set the covariance matrix initial value P_0, forgetting factor λ and also model's order by configuring the dimension of the regression vector, respectively by setting n_y and n_u, as described previously on section 4.2.1.

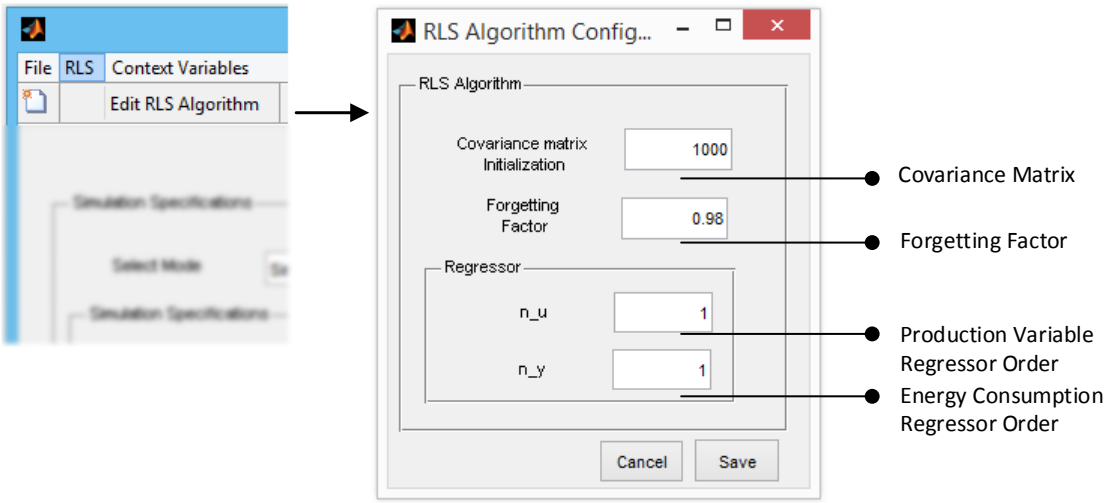

Figure 5.14: RLS algorithm specifications.

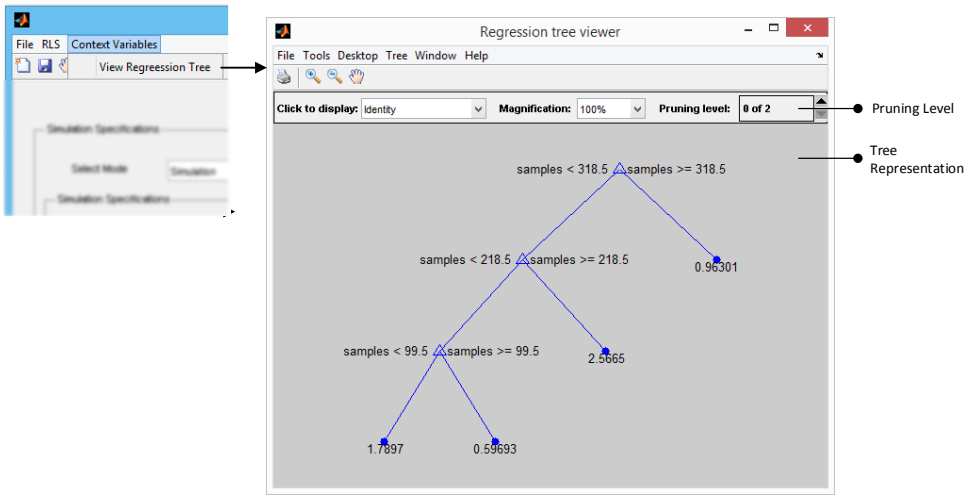

Figure 5.15: Regression tree viewer.

The context variable menu, allows the user to view the tree representation of the regression tree algorithm, as depicted on figure 5.15. Also, the user is able to manipulate the number of branches and leafs by pruning the tree.

The use case UML diagrams for application-specific tools is shown on figure 5.16 (A) and RLS specifications and regression tree view is depicted on (B).

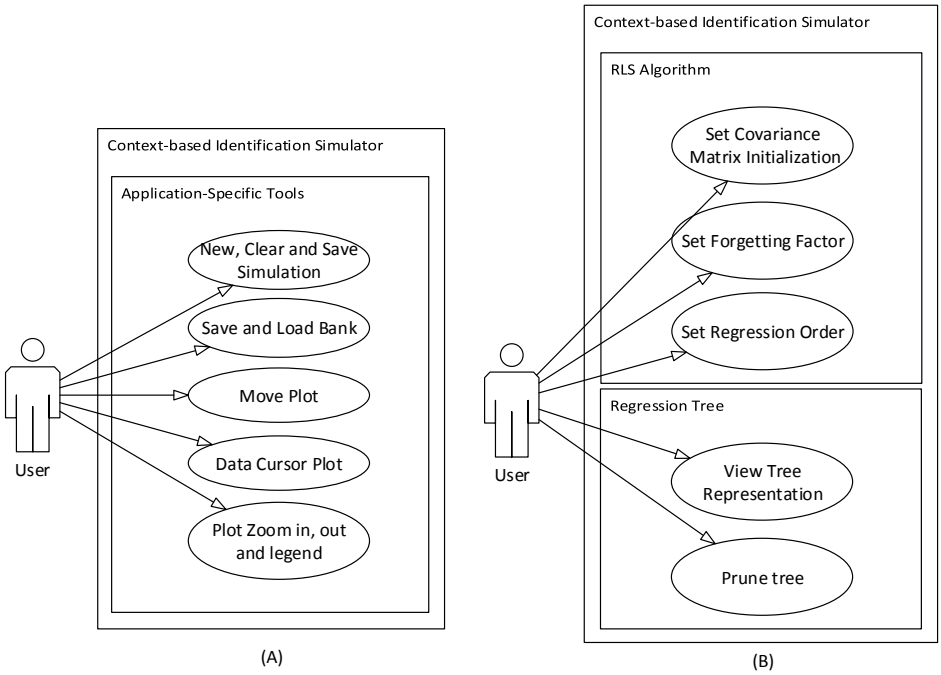

Figure 5.16: Use case UML diagrams. (A) Application-specific tools. (B) RLS algorithm specifications and regression tree.

6

Experimental Results

In this chapter the developed CBEM system is applied to a real cement plant milling unit.

A brief overview of the cement manufacturing process and in particular the cement grinding circuit is depicted. Also, the acquired raw data is presented and analysed.

Finally, results from the CBEM system performance are shown and detailed.

6.1 Cement Manufacturing Process and Grinding Circuit

The cement manufacturing process is normally divided into two large steps. The first step is the clinker production from raw material and the second step is the cement production from clinker.

Raw materials are normally quarried and delivered in bulks, hence they have to be crushed on raw material milling units, and then stored and homogenized on proper silos. This mixture is then preheated and fed to a rotary kiln, where chemical reactions between materials at high temperature take place, producing the so called clinker [64].

The second stage is handled on a cement grinding circuit. Before the grinding phase, several materials such as gypsum, limestone, tras and other additives are added to the clinker. At this stage, the type of commercial cement under production is defined, which is characterized by the ratio of clinker on the total mixture. All constituents are ground on cement milling units, leading to a fine and homogeneous powder called cement. Lastly, cement is then stored in silos before being dispatched to the market.

Figure 6.1 depicts a typical cement manufacturing process.

Of the cement production chain steps, grinding and milling operations are rather energy inefficient. As depicted on chapter 2, the primary energy consuming units are the

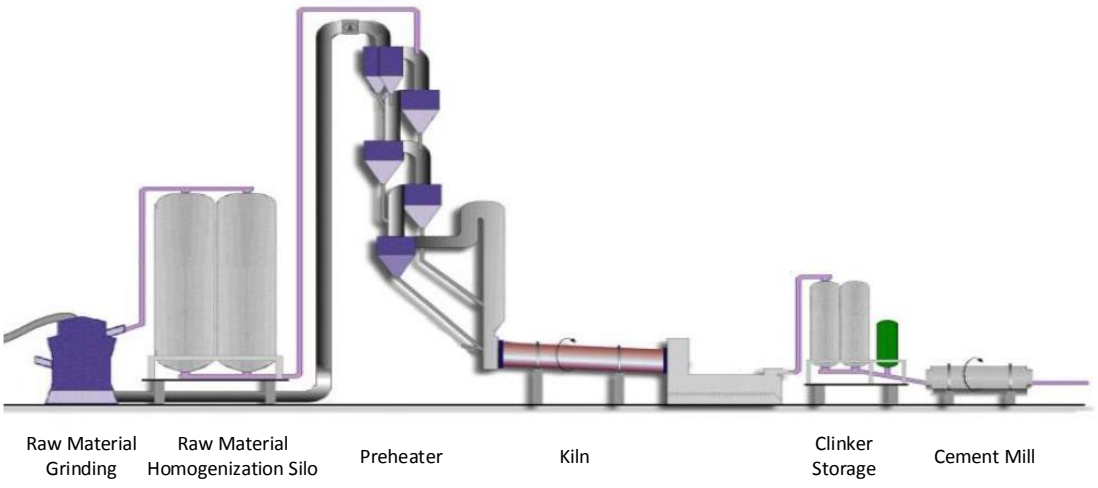

Figure 6.1: Typical cement plant manufacturing process. Adapted from [65].

cement mills, consuming about 38% of the overal plant consumption, which drives our focus on to the grinding circuit.

The cement grinding circuit is normally composed by a multi-compartment tube mill, separated from each other by a slotted diaphragm, in closed circuit with an air separator as illustrated on figure 6.2 [6].

According to the type of cement under production, the mixture of clinker with the other components may produce a fresh feed of material to the cement mill that is harder or easier to grind. Hence, as the ground material passes through the air separator, particles are selected to be recirculated in the grinding circuit in order to achieve the desirable level of narrowness, consuming more energy in this process.

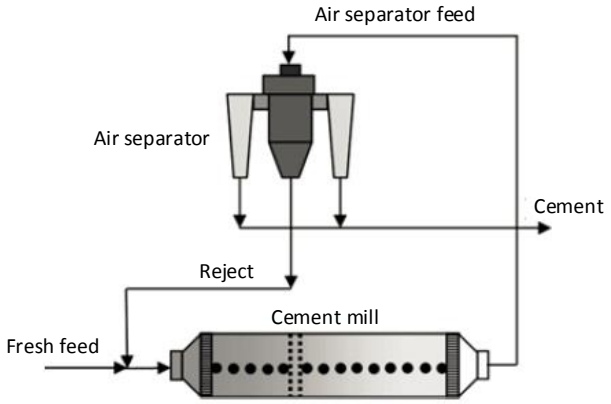

Figure 6.2: Cement grinding circuit structure. Adapted from [6].

.

6.2 Cement Mill Data

Data from a single cement milling unit was collected during 5 months of operation, from September 2012 to January 2013, by the cement plant own monitoring system, with a sampling period of 10 minutes. The set of collected data comprises measurements of the cement mill input materials: flow of clinker F_c; gypsum F_g; limestone F_l; trass F_t and the air separator output: flow of finished cement F_{ce}. All flows are measured in tonnes per hour (t/h). Measurements of the power supplied P_{drive} in MW were also taken.

Due to the large amount of information and for ease of analysis, the months of September 2012, October 2012 and January 2013 which depict the production of different types of cement are analysed in detail, although the same analysis for the months of November 2012 and December 2012 is performed and depicted in appendix A.1.

Figures 6.3, 6.4, 6.5 depict the raw data obtained from the cement mill operation, which comprises material flows in (A) and power consumption in (B), respectively for the months of September, October and January.

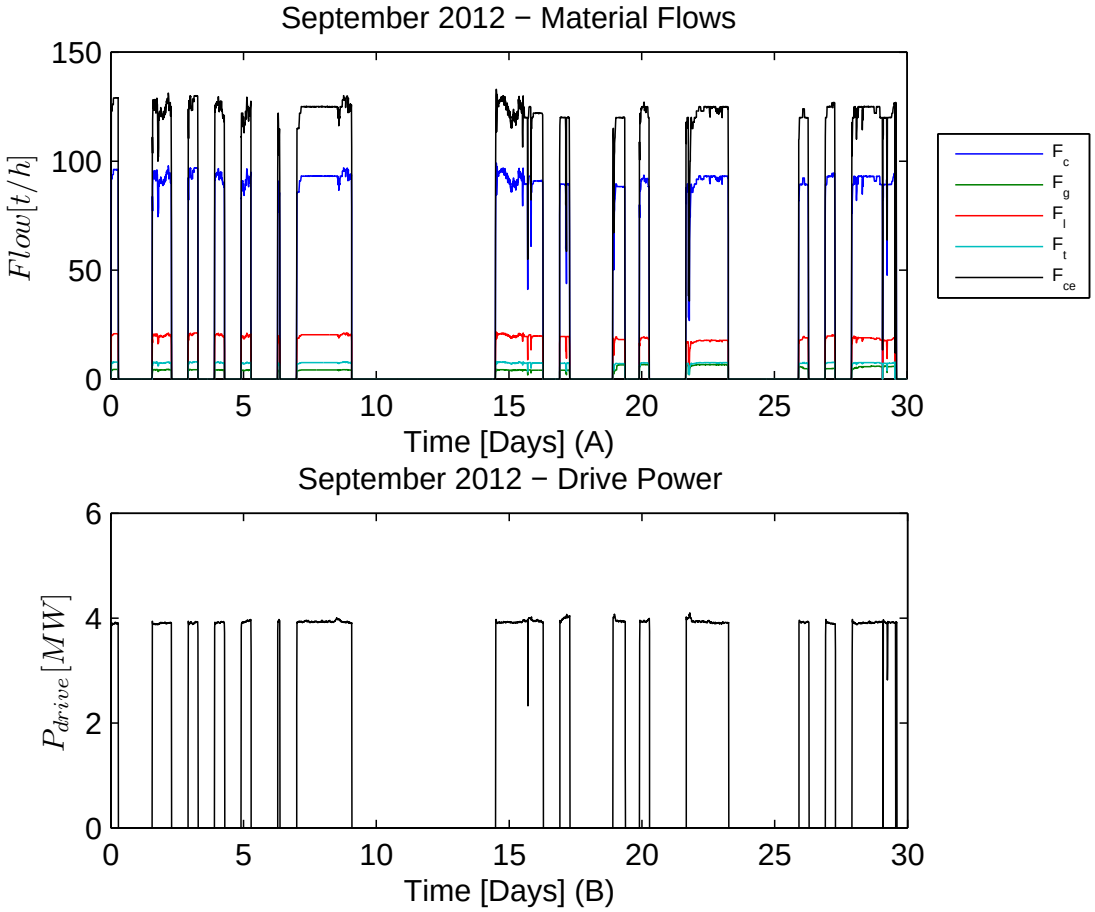

Figure 6.3: September 2012 raw data. (A) Flow of input and output materials. (B) Power consumption.

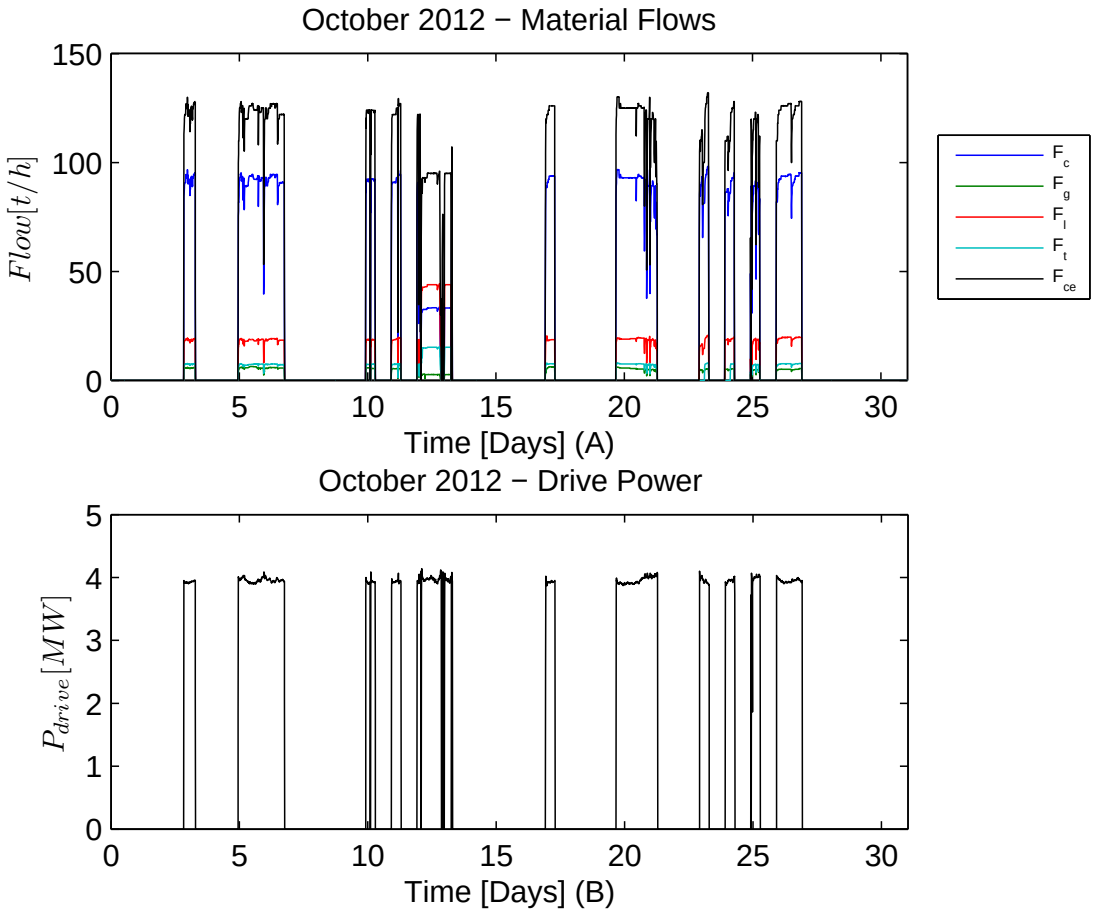

Figure 6.4: October 2012 raw data. (A) Flow of input and output materials. (B) Power consumption.

During September, the cement flow and the materials mixture type are nearly constant over the entire month. Approximately a total of 33.2 kt of cement were produced with an average flow rate of 121.6 t/h. The power consumption is nearly constant, rounding about 4 MW during the full month, in which the mill operated approximately 270 hours, where the minimum operation time was of 2 hours.

Regarding the month of October, the materials mixture is different for a specific set of days between the 12^{th} and the 14^{th}. During this period a total of 2.3 kt of cement were produced with an average production rate of 89.7 t/h, whereas in the remaining month, a total of 20.8 kt of cement was produced with an average production rate of 119.5 t/h, similar to the production rate shown through the entire month of September. Also, the minimum production time was of 40 minutes out of over 200 hours of production. Another important fact is that when the milling unit is operating, the power consumption is approximately the same, despite the different cement production rate between the 12^{th} and the 14^{th} days.

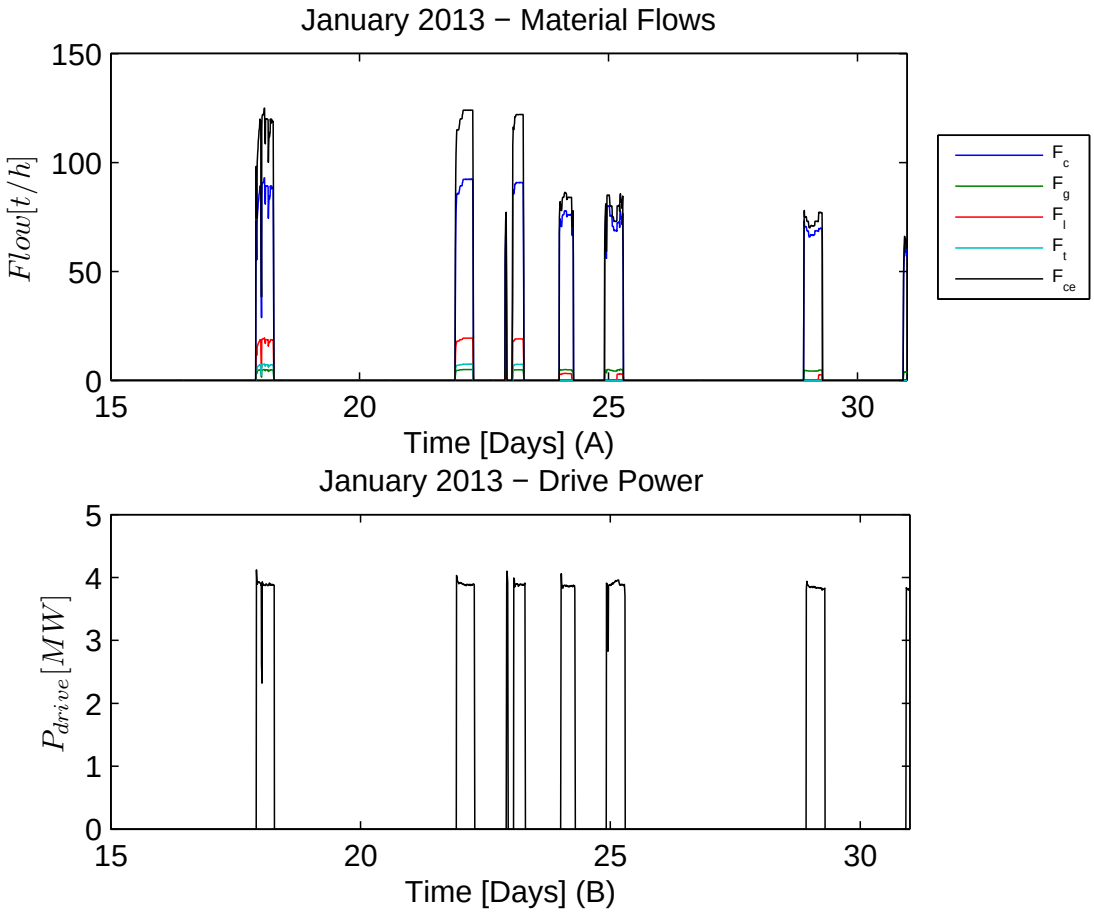

Figure 6.5: January 2013 raw data. (A) Flow of input and output materials. (B) Power consumption.

During January, the cement mill was inactive for more than half of the month. However, it is interesting to analyse the data from the 24th day until the end of the month. During this period, cement was produced with an average production rate of 78.1 t/h, different from the production rates shown on the previous months. The mixture of materials in this period is different as well. On the remaining month, the production rate was of 118.9 t/h, similar to the full month of September and most of October.

Clearly, the cement production rate varies with the materials mixture types that are being milled. As described previously, the amount of clinker in the total material fed to the cement mill, defines which type of cement is under production. Moreover, depending on the clinker ratio, the resultant mixture is a cement that is easier or harder to mill, which may have to be recirculated in the grinding circuit in order to achieve a desirable level of particle size, consuming more energy in this process. While the cement mill is operating, independently of which type of cement being produced, the power drive uses the same amount of power, nearly 4 MW. However, as the cement might have to be re-circulated, a MWh of energy does not produce the same amount of cement of each type.

Hence, from a energy consumption point of view, an interesting approach is to evaluate for each type of cement how many tonnes are produced with a single MWh of power. This performance indicator is denominated as specific consumption y_{sc} and is define as follows:

$$y_{sc}(t) = \frac{P_{drive}(t)}{F_{ce}(t)} \tag{6.1}$$

Where, P_{drive} is the power drive power in MW and F_{ce} is the cement flow.

The cement type is defined by the clinker ratio Q in the total material , described as follows:

$$Q(t) = \frac{F_c(t)}{F_{total}(t)} \tag{6.2}$$

Where F_{total} is the sum of all individual material flows entering the mill:

$$F_{total}(t) = F_c(t) + F_g(t) + F_l(t) + F_t(t). \tag{6.3}$$

F_c, F_g, F_l, F_t are respectively clinker, gypsum, limestone and trass flows.

Figures 6.6, 6.7 and 6.8 depict the specific consumption in (A) and the cement type produced in (B), respectively for the months of September, October and January.

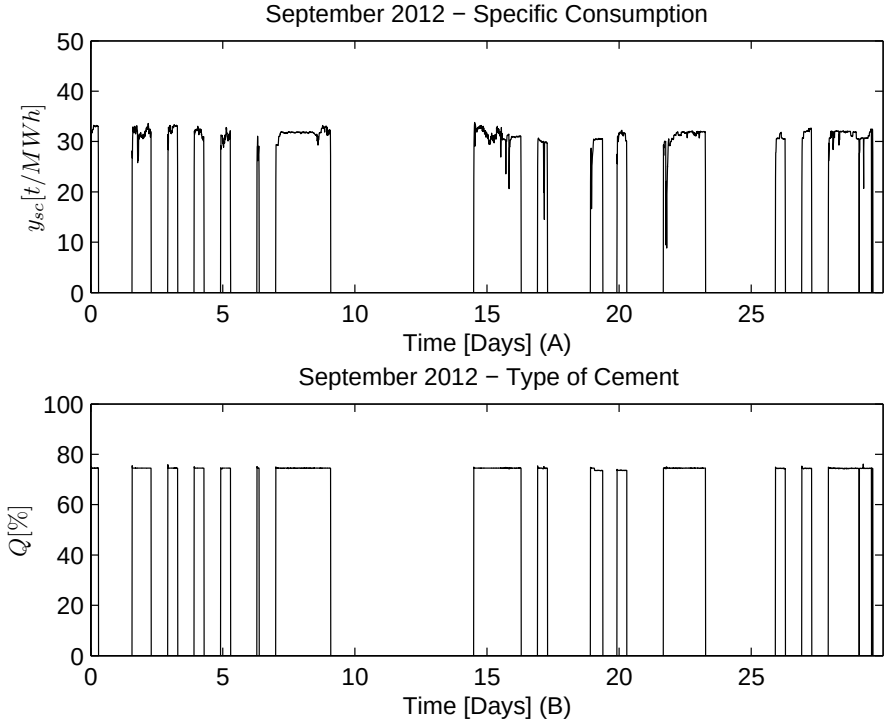

Figure 6.6: September 2012 - Cement specific consumption and cement type. (A) Specific consumption. (B) Cement type.

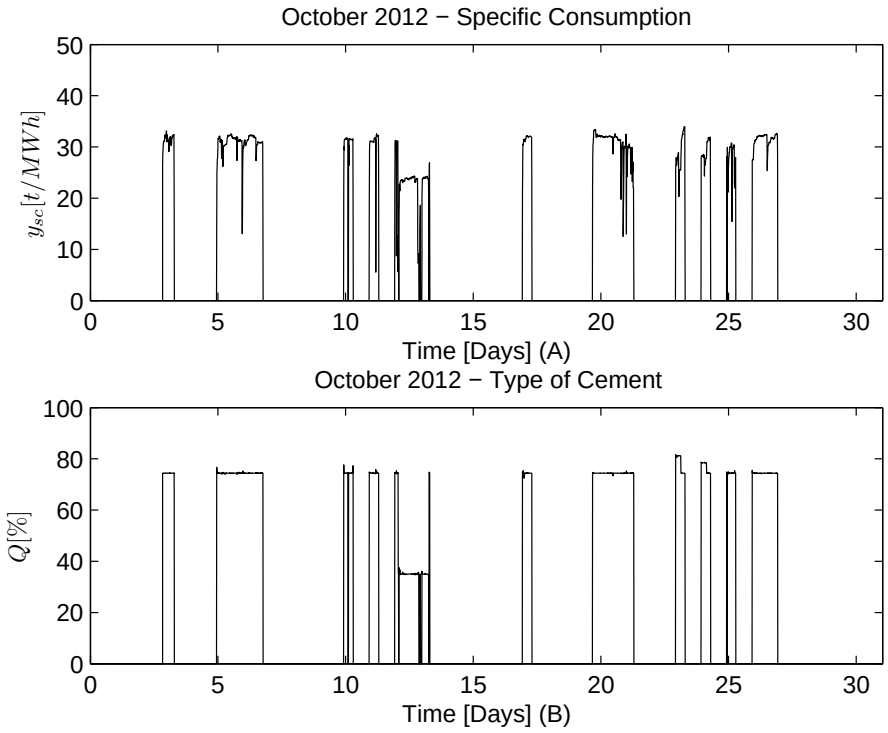

Figure 6.7: October 2012 - Cement specific consumption and cement type. (A) Specific consumption. (B) Cement type.

During September month, the clinker ratio in the total material flow is nearly the same, rounding about 75%, indicating that the same type of cement is being produced. When the mill is operating, the specific consumption is of 31 t/MWh, this means that with a single MWh, 31 tonnes of cement are produced. For ease of reference, consider that during September the cement produced is of type A, defined by a ratio of clinker of 75%.

During October month, more specifically from the beginning of the 12^{th} day until the 14^{th}, the clinker ratio in the total material flow is clearly different, rounding about 35%. This indicates that a different type of cement is being produced, lets denominate it as of type B, defined by a clinker ratio of 35%. The specific consumption for this period is nearly 23 t/MWh, indicating that for type B cement production, a single MWh produces about 24 tonnes of final product. On the remaining month, the type of cement produced is of type A.

Regarding January, after the 24^{th} day, a new type of cement is produced, defined by a clinker ratio of approximately 95%, lets name it as of type C. In this period, the specific consumption is nearly 20 t/MWh, hence a MWh produces 20 tonnes of type C cement. On the rest of the month, type A cement was produced.

Therefore, with a single MWh, significant different quantities of final product are produced. For type A cement, a MWh produces 31 tonnes, whereas for cement type B and C, it produces 24 and 20 tonnes, respectively. From an energy consumption point

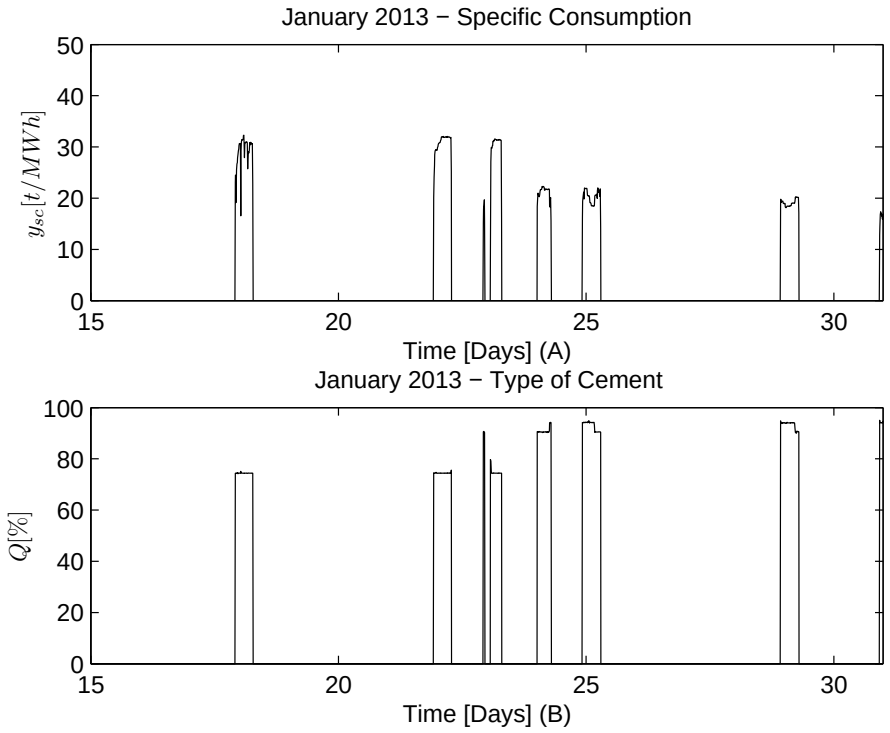

Figure 6.8: January 2013 - Cement specific consumption and cement type. (A) Specific consumption. (B) Type of cement.

of view, this means that for producing the same of amount of final product for each type of cement, more energy is consumed for types B and C than for type A. Table 6.1 shows the data summary focusing the most relevant information for the 3 months operation.

Automatically identifying these operating contexts enables a more accurate energy consumption forecasting. Furthermore, it provides a tool for analysing how the contexts impact energy consumption. The CBEM approach is applied to this test case and detailed results are described in the following chapters.

Table 6.1: Data Summary for the months of September, October and January.

Data	September 2012	October 2012	January 2013
Type of Cement	A	A ; B	A ; C
Clinker Ratio [%]	75	75 ; 35	75 ; 95
Production Rate [t/h]	121.6	119.5 ; 89.7	118.9 ; 78.1
Total Production [kt]	33.02	23.14	4.73
Specific Consumption [t/MWh]	31.03	30.43 ; 22.52	29.4 ; 19.8
Shortest Operating Time [min]	120	40	60

6.3 Context-based Energy Models Results

6.3.1 CBEM Initialization

The CBEM overall goal is to be able to accurately estimate specific consumption $\hat{y}_{sc}$ for each cement type. The specific consumption provides significant information regarding how many tonnes of cement are produced with a single MWh, which is a very important performance indicator as it allows the forecasting of the energy demand to achieve a desirable production rate.

Clearly, the type of cement under production directly affects the energy consumption of the milling circuit, hence from a CBEM point of view, the type of cement is considered to be a context variable, which is given by the ratio of clinker on the total material fed to the cement mill. As the acquired data has the flows of material measured separately, the type of cement is obtained by dividing the flow of clinker F_c by the sum of all individual material flows F_t. Figure 6.9 depicts the CBEM representation of the milling unit, along with the quantities at stake.

During the entire set of data, 3 different types of cement were produced, characterized by a clinker ratio on the total mixture of 75%, 35% and 95%, respectively for cement types A, B and C. Hence, in terms of system initialization, the bank sensitivity value to context changes γ is set to 15% in order to be able to distinguish between cement types. Furthermore, it was shown that the minimum operating time was of 40 minutes, hence the context variable period of steadiness τ is set to 40 minutes, in order to properly partition the type of cement per operating regime. The centroid tolerance κ is set according the default value of 10^{-6}. In terms of RLS initialization, a second order model as shown good results in [66]. The forgetting factor is initialized with 0.99.

Table 6.2 shows the initialized values.

The analysis of the CBEM system comprises 2 main points for each month under analysis. The first is the system performance in identifying when a type of cement is being produced and automatically create or select a local model to estimate the specific

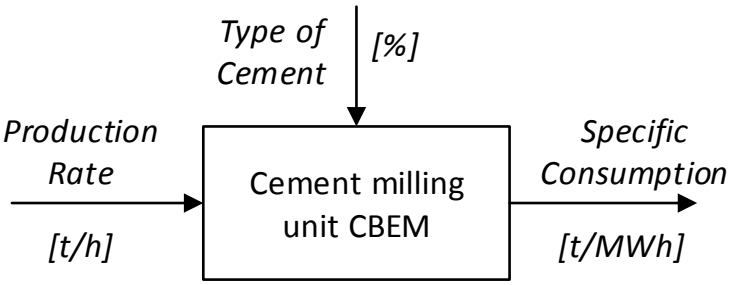

Figure 6.9: CBEM representation of the milling unit.

Table 6.2: CBEM system initialization values.

γ	τ	κ	λ	Θ
15 [%]	40 [min]	10^{-6}	0.99	$[a_1\ a_2\ b_1]$

consumption. This is performed by partition the ratio of clinker in the total mixture into operating contexts and calculating the respective centroid for each one. Afterwards, according to the centroid of each operating context, a model is selected. The second point is the specific consumption estimation for each type of cement.

For ease of analysis, results are organized by months: September 2012, October 2012 and January 2013. The analysis performed on the months of November 2012 and December 2012 are described on the appendix A.2.

6.3.2 CBEM Results of September 2012

During the month of September, only type A cement was produced. Figure 6.10 depicts the regression tree algorithm and centroid calculation for the entire month in (A), along with the model selected to be active for estimation in (B), for a global view of the system performance in terms of context identification.

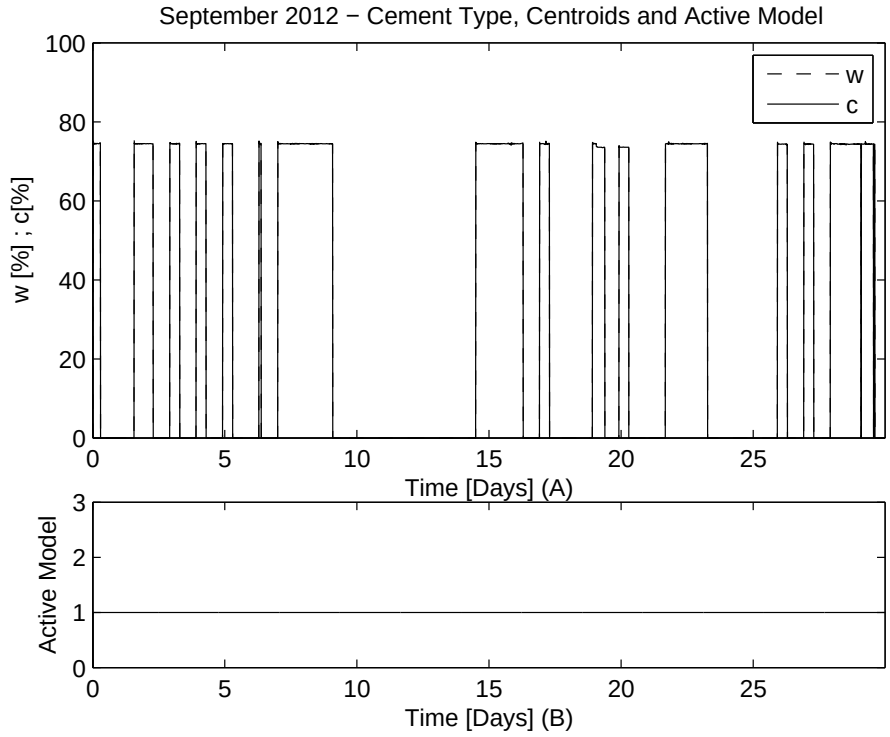

Figure 6.10: CBEM context identification results of September 2012 - (A) Type of cement and centroids zoom on the full month. (B) Active model.

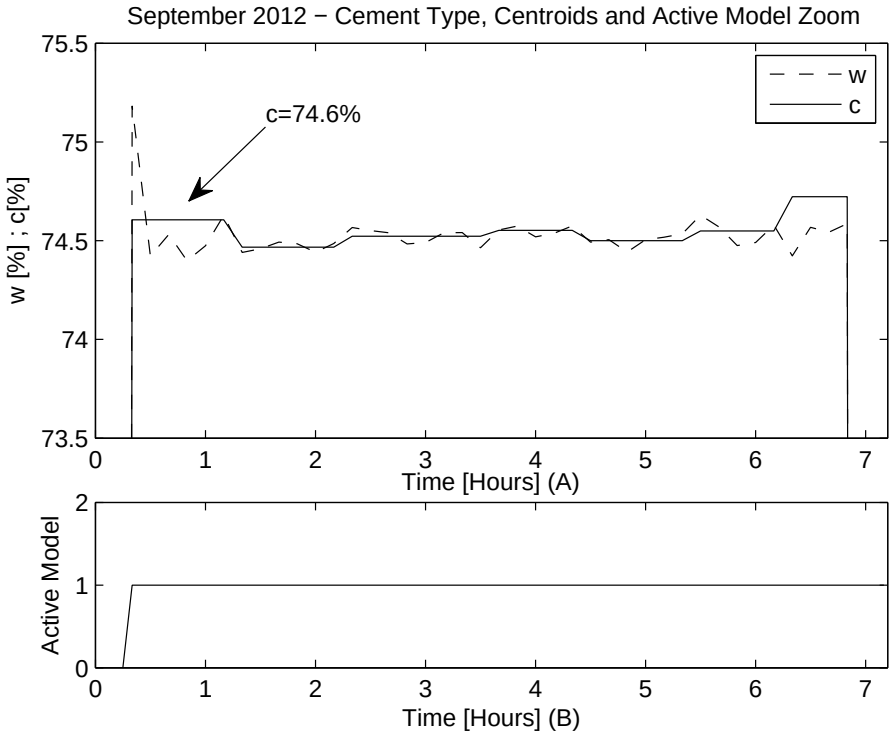

Figure 6.11: CBEM context identification results of September 2012 - (A) Type of cement and centroids zoom on first cement type A production. (B) Active model.

Right at the beginning of the month, the type A cement is automatically identified. For a better understanding of the system behaviour, figure 6.11 depicts a zoom of the previous figure on the first operating regime where cement type A is produced.

From figure 6.11, it is shown that the regression tree algorithm is able to partition the first cement type A production time into several regions and calculate centroids for each one. As the production was performed continuously during about 7 hours and the regression tree algorithm was initialized with a period of steadiness of 40 minutes, several partitions are made. After the first centroid region is calculated ($c = 74.6\%$) a search in the bank of models is performed in order to retrieve a valid model for the current centroid. As the bank is empty, a new model is created, initialized with the centroid of the current region and selected to be active for estimation. Same logic applies to the second centroid calculated, however in this case as the centroid calculated falls within the current model validity limits, instead of creating a new one, the previous is maintain as active to continue estimation. Same happens for the rest of the month, whenever the cement mill is operating.

Figure 6.12 depicts the prediction of the CBEM system and compares it to the real specific consumption in (A) and shows the evolution of the model parameters estimation in (B). For a better visualization of the CBEM prediction accuracy, a zoom is made over a single day of operation in (C).

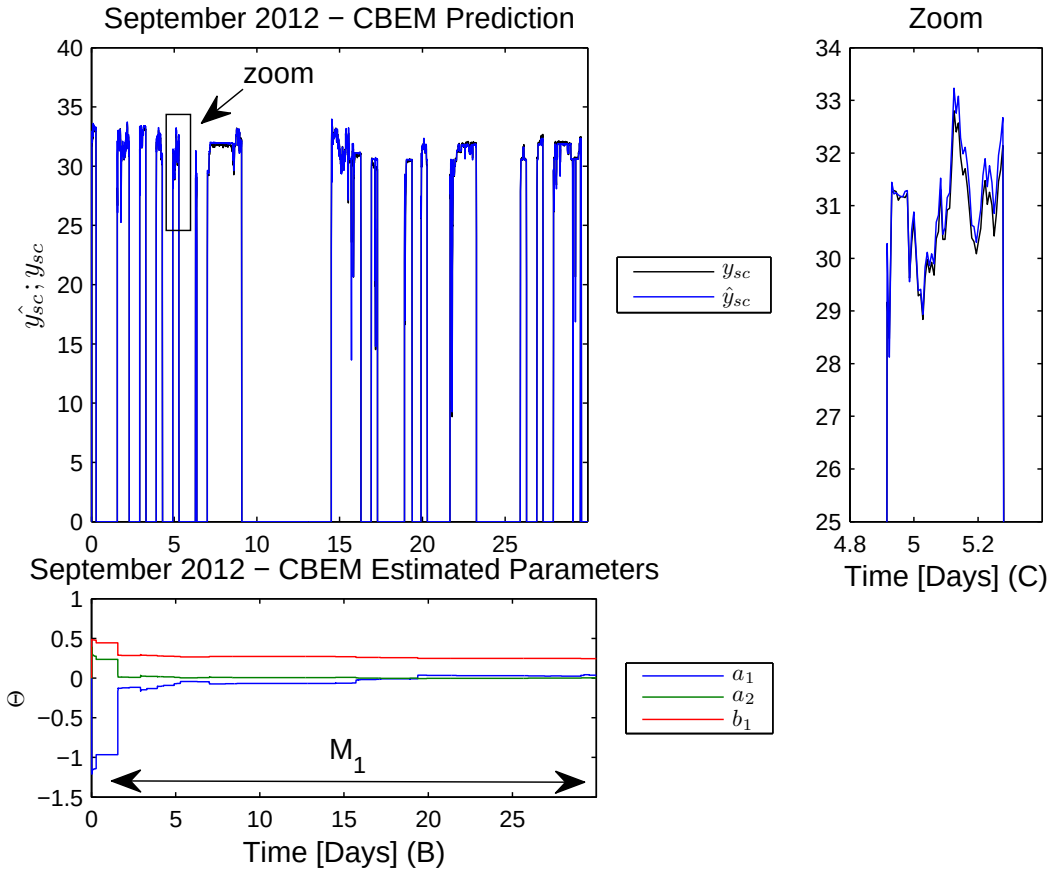

Figure 6.12: CBEM prediction results of September 2012 - (A) Specific consumption prediction. (B) Estimated Parameters. (C) Prediction over a day of operation.

The CBEM estimation shows a very accurate prediction with an average estimation error for the entire month lower than 0.09 t/MWh. Also, the parameters estimated are able to converge to a nearly constant value, indicating a steady estimation. Table 6.3 shows the bank of models, respective centroids and parameters, for the month of September. As only type A cement was produced, only one model was active during the entire month. Consumption estimation and estimation error are shown on table 6.4.

Table 6.3: Models Estimated in September 2012.

Cement Type	**Model**	**Centroid [%]**	**Estimated Parameters**
A	M_1	$c_{M_1} = 74.61$	$\Theta_{M_1} = [0.0372\,;\ 0.0038\,;\ 0.2446]$

Table 6.4: CBEM average absolute estimation error for September 2012.

CBEM Average Absolute Estimation Error $\overline{\lvert\Delta\epsilon\rvert}$
$\overline{\lvert\Delta\epsilon\rvert} = 0.085\, t/MWh$

6.3.3 CBEM Results of October 2012

During the month of October, a new type of cement was produced: type B. Figure 6.13 shows the context identification for the entire month of October in (A), highlighting the days where type B cement was under production.

As the centroid of the region corresponding to type B is of 35%, it exceeds the validity limits of model M_1, hence a new model M_2 is automatically initialized and selected to be active for estimation as depicted in (B).

On the rest of the month, cement type A is produced, hence model M_1 is again selected to be active, loading its covariance matrix and previous estimated parameters to continue estimation of cement type A influence on the specific consumption.

Regarding the months of November and December the same types of cement were produced, hence it is not of special interest detailing the system behaviour since it is very similar to this month. However, a similar although not so detailed analysis can be consulted on appendix A.2.

Concerning consumption estimation, figure 6.14 depicts the CBEM prediction compared with the real cement mill specific consumption in (A) and the estimated parameters in (B).

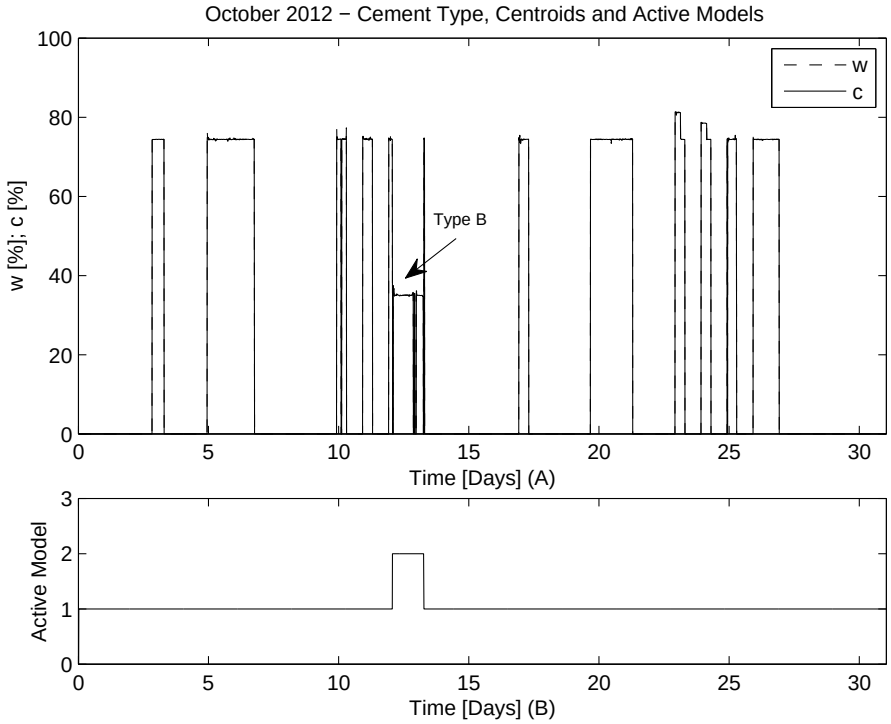

Figure 6.13: CBEM context identification results of October 2012 - (A) Type of cement and centroids. (B) Active model.

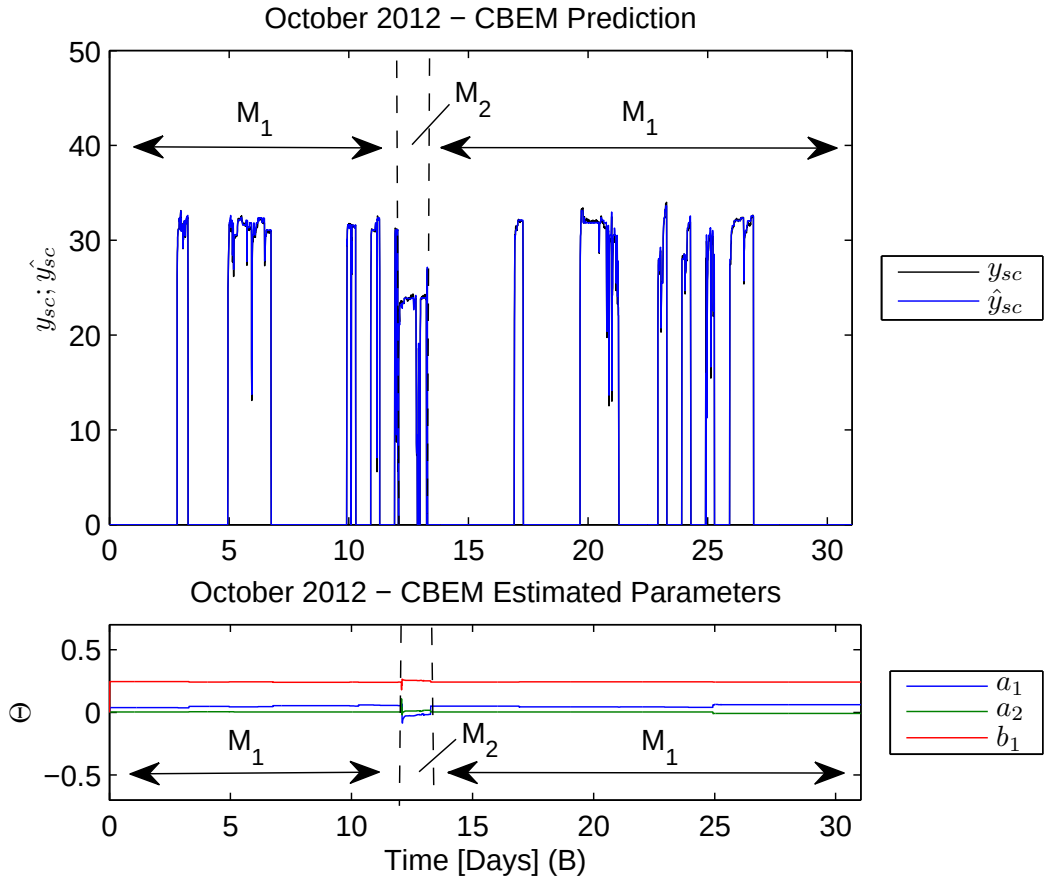

Figure 6.14: CBEM prediction results of October 2012 - (A) Specific consumption prediction (B) Estimated Parameters.

As depicted, the specific consumption is considerably different for each type of cement. However, the CBEM prediction remains very accurate as a new model was automatically created to estimate specific consumption specifically for the region where cement type B was being produced. Figure 6.15 depicts a zoom on the plant consumption and CBEM prediction when the operating context changes from type A cement to type B in order to best depict the impact of context on the consumption. As it is shown, although specific consumption changes, the CBEM prediction remains accurate for both types of cement due to the commutation between model M_1 and M_2.

The average absolute estimation error for the entire month is again lower than 0.09 t/MWh, similar to the estimation of September. Such an accuracy is achieved due to the context identification capability of the CBEM system. Traditional energy consumption models would largely fail in this situation as demonstrated theoretically, due to the use of the same parameters trained with the full data, disregarding context awareness. Furthermore, the model prediction when type A cement is under production would also be affected, due to the influence of type B cement in the overall parameter estimation.

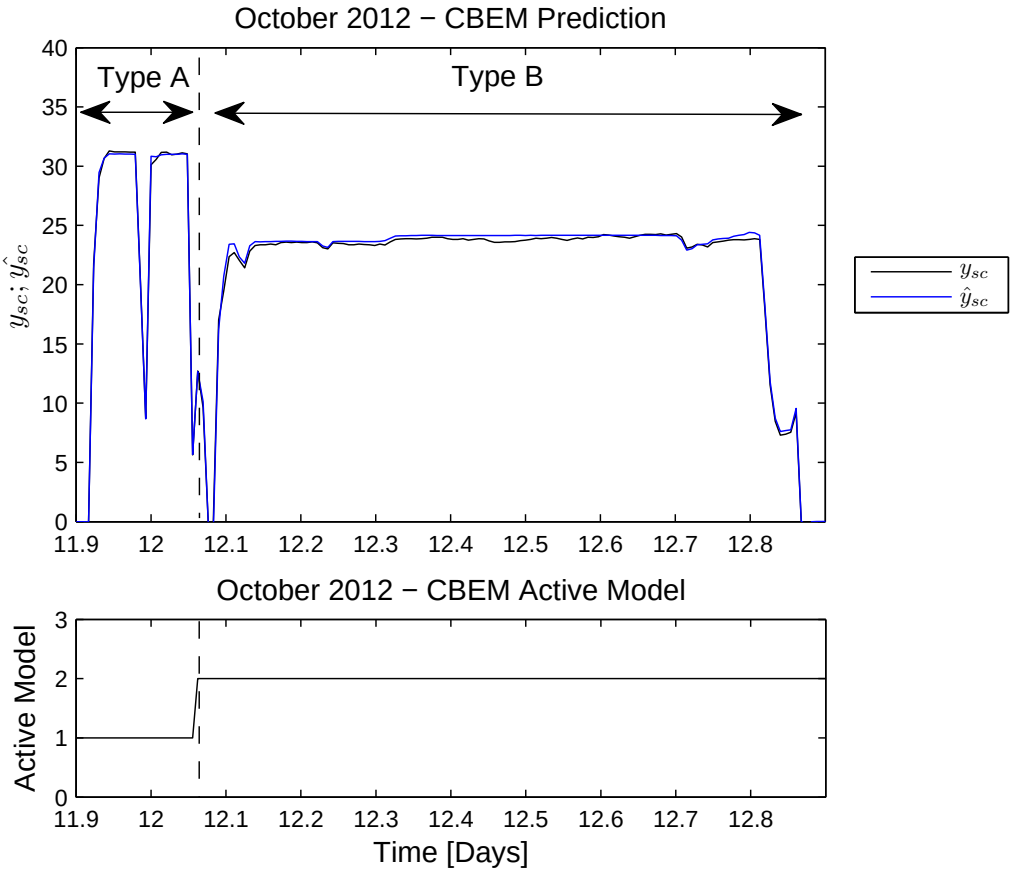

Figure 6.15: CBEM prediction results of October 2012 zoom on context commutation - (A) Specific consumption prediction (B) Active model.

During this month, a new model was created for type B cement, but also the existing model for type A was optimized, allowing a continuous update of the model parameters hence improving estimation.

Table 6.5 shows the models estimated until the end of this month, respective centroids and parameters. Prediction results are shown on table 6.6.

Table 6.5: Models Estimated in October 2012.

Cement Type	**Model**	**Centroid [%]**	**Estimated Parameters**
A	M_1	$c_{M_1} = 74.61$	$\Theta_{M_1} = [0.0612\ ;\ 0.0086\ ;\ 0.2411]$
B	M_2	$c_{M_2} = 35.59$	$\Theta_{M_2} = [0.0174\ ;\ 0.0171\ ;\ 0.2519]$

Table 6.6: CBEM average absolute estimation error for October 2012.

CBEM Average Absolute Estimation Error $\overline{\lvert\Delta\epsilon\rvert}$
$\overline{\lvert\Delta\epsilon\rvert} = 0.081\ t/MWh$

6.3.4 CBEM Results of January 2013

Regarding January 2013, cement production only starts again on the last half of the month, where cement type A and a new cement type is produced. Lets name this new type as type C.

This type of cement is identified by the regression tree as having a centroid of approximately 95% as depicted in figure 6.16 (A) while the other cement types have centroids of approximately 75% and 35%. As the CBEM system was initialized with a bank sensitivity of 15%, non of the existing models validity limits contain the centroid of type C. Thus, a new model M_3 is automatically created and selected to be active for estimation during type C production time as depicted in (B).

Regarding the CBEM system consumption estimation, it is very accurate for this month as depicted in figure 6.17 (A) and specially zoomed on the type C production regime in (C).

The average absolute estimation error is of nearly 0.02 t/MWh, which depicts an improvement when compared with the previous months, due to the continuous optimization and learning of the estimated parameters over time. Also, in figure 6.17 B it is observed a variation on the estimated parameters when type C cement is being produced. This is due to the impact that this type of cement has on the cement mill dynamic, which is continuously captured by model M_3.

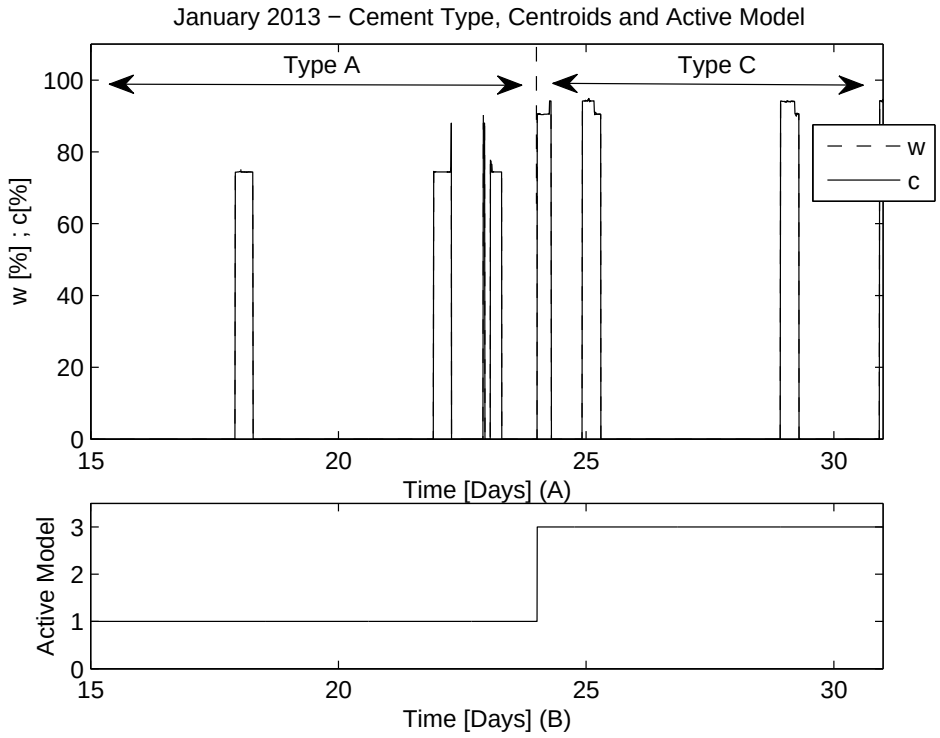

Figure 6.16: CBEM context identification results of January 2013 - (A) Type of cement and centroids. (B) Active model.

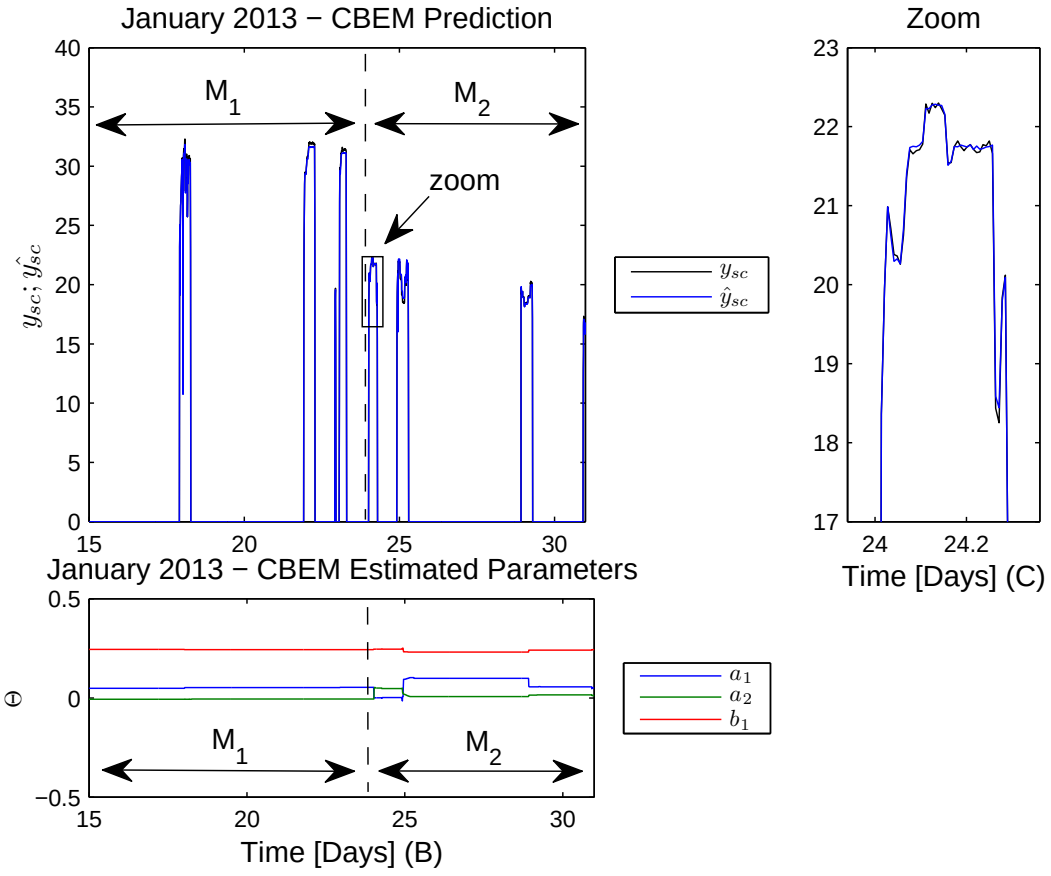

Figure 6.17: CBEM prediction results of January 2013 - (A) Specific consumption prediction. (B) Estimated Parameters. (C) Zoom on a type C operating regime.

During this month, cement type A and C were produced, thus model M_1 from the previous months was selected to be active and a new model M_3 was initialized for estimation when type C cement was under production. Table 6.7 shows the models estimated and table 6.8 the average absolute estimation error for this month.

Table 6.7: Models Estimated in January 2013.

Cement Type	Model	Centroid [%]	Estimated Parameters
A	M_1	$c_{M_1} = 74.61$	$\Theta_{M_1} = [0.0452\ ;\ 0.0060\ ;\ 0.2427]$
C	M_3	$c_{M_3} = 94.51$	$\Theta_{M_3} = [0.0511\ ;\ 0.0116\ ;\ 0.2432]$

Table 6.8: CBEM average absolute estimation error for January 2013.

CBEM Average Absolute Estimation Error $\overline{\lvert\Delta\epsilon\rvert}$
$\overline{\lvert\Delta\epsilon\rvert} = 0.022\, t/MWh$

6.4 Results Summary and Discussion

During the 5 months of cement mill operation, the CBEM system was able to automatically identify and distinguish between operating contexts, by identifying when different types of cement were under production.

For each type of cement, a specific model was selected to be active for estimation, capturing the influence that the type of cement under production has on the cement mill consumption. The resultant bank of models after the 5 month operation is shown on table 6.9, along with the respective centroids and estimated parameters. Parameters a_1 and a_2, respectively for the regressor samples of $y_{sc}(t_k - 1)$ and $y_{sc}(t_k - 2)$ show a significant difference between each model, which is an accordance to the fact that different types of cement change the cement mill dynamic in terms of consumption, due to the recirculation of material in the cement grinding process.

Each cement type significantly impact the specific consumption of the cement mill operation, leaving distinct energy consumption footprints, which were successfully captured by the CBEM system by dynamically creating new energy models, continuously optimized over time for each type of cement. The continuous optimization is noticed by the decrease of the estimation error as table 6.10 shows. The overall estimation accuracy of the CBEM system is proven, by an average absolute estimation errors lower than 0.08 t/MWh.

The identification of context-based energy models, allows the company an understanding about how energy is used across the production chain. Particularly, in this test case, it was observed and quantified that a MWh of power produces a significantly higher amount of type A cement than of type B or C, showing that the cement mill is more energy efficient for cement type A production.

As the energy footprint was capture individually for each type of cement, it is possible, for instance, to forecast how much energy will be needed (thus how much it will cost) to achieve a desirable production goal. This information when join together with the

Table 6.9: Bank of models after 5 months operation.

Cement Type	Model	Centroid [%]	Estimated Parameters
A	M_1	$c_{M_1} = 74.61$	$\Theta_{M_1} = [0.0452\ ;\ 0.0009\ ;\ 0.2488]$
B	M_2	$c_{M_2} = 35.59$	$\Theta_{M_2} = [0.0312\ ;\ 0.0059\ ;\ 0.2416]$
C	M_3	$c_{M_3} = 94.51$	$\Theta_{M_3} = [0.0511\ ;\ 0.0116\ ;\ 0.2432]$

Table 6.10: Summary of estimation results.

Month	Sept 2012	Oct 2012	Nov 2012	Dec 2012	Jan 2013
Cement Type	A	$A\ ;\ B$	$A\ ;\ B$	$A\ ;\ B$	$A\ ;\ C$
Active Models	M_1	$M_1\ ;\ M_2$	$M_1\ ;\ M_2$	$M_1\ ;\ M_2$	$M_1\ ;\ M_3$
$\overline{\lvert\Delta\epsilon\rvert}$ **[t/MWh]**	0.085	0.081	0.088	0.075	0.022

company's production strategy, offers valuable decision making support. As an example, using the context-based energy consumption models, simulations can be performed in order to evaluate if a new milling unit, more suitable for cement types B and C production, should be bought. Moreover, an estimation of how much time would it take to recover the investment, just by energy savings may also be derived. Other analysis may be performed using the estimated models, as the calculation of the maximum value, allowing the identification of consumption peaks.

As a summary, during the 5 month operation, 3 types of cement production were identified and 3 local energy models were estimated, whose estimation errors showed to be very accurate. The incorporation of the operating context into the model's structure allows a direct mapping between the cause (in this case the type of cement, represented by its centroid) and the effect (model's estimation of specific consumption). Both regression tree used for context identification and RLS for consumption estimation proved to be a very good approach for building context-based energy consumption models.

7

Conclusion and Future Work

7.1 Conclusion

Given the increasing trend of energy consumption and the high necessity to find new and alternative ways to reduce consumption worldwide and in particular in the industrial sector, the main goal of this dissertation is to propose and develop a new approach for identifying possible scenarios of consumption reduction by tracking down how energy is used in industrial plants. As energy consumption in industrial environments is influenced by several factors, this dissertation proposes a method for automatically identifying the context in which the plant/process is operating in order to estimate energy consumption models that are able to capture context influence on the plant consumption, thus mapping consumption causes to consumption patterns.

Regression tree algorithms proved to be a suitable method for identifying context changes. The ability to handle both continuous and categorical variables is a very important feature in industrial plants, as it not only allows an extension of variables type that can be used for context identification as it also provides an easier support for decision-making processes which often rely on categorical values. The tree look-a-like representation derived from the regression tree algorithm is also an important feature to understand, pinpoint and explain where and why consumption patterns are related with context variables.

The context variables partition into regions which determine an operating context, each one identified by a unique centroid, provides a natural structure to support a multi-model approach, where each local model estimates consumption for a certain operating context. This approach, allows the mapping between a certain influencing factor (the

cause), and energy consumption pattern (the effect), resolving one of the existing problems of the current state of the art methodologies: the lack of structure to incorporate context influence for identifying causes of consumption. Also, the incorporation of the operating context centroid in a multi-model bank architecture to automatically identify if a new model should be created or an existing one selected for estimation, not only provides a solid structure to relate context with consumption, but also allows the system to dynamically carry out energy consumption estimation and a continuous and optimized learning of each local model, hence improving prediction quality.

Regarding the proposed approach for handling multiple context variables, throughout the dissertation and particularly noticed during the experimental application, it was realized that if a single context variable is well chosen, there is no significant gain in adding additional complexity over the system. Hence, the dissertation was more focused towards a single context variable influence than of multiple, although an approach was proposed, which may lead to future work investigation.

The simulator developed to best understand and test the theoretical formulation of the proposed approach, allows the configuration of the overall system, plant behaviour and even tuning for application on real plants, as it allows the loading of real measurements. The simulator was developed to be user friendly and to provide graphical and numerical results, which makes it not only a good tool to understand the system, but also to evaluate how the system would perform if applied to a specific plant on an implementation case. Alongside the development of this simulator, other relevant work on the fields of computational simulation of dynamical systems and control systems, which resulted in 5 publications on 2 international conferences and 1 journal (consult [67, 68, 69, 70, 71]), was in a sense, drove from the simulator developed in this dissertation.

Lastly, the context-based energy consumption system was applied to a real cement grinding circuit of a cement plant, with focus on the cement mill operation. The developed system ran over the information collected during 5 months and showed excellent results that prove the applicability of the proposed approach on industrial plants. During the cement mill operation, the system was to able to identify 3 different operating contexts based on a context variable defined by the type of cement that was under production. During the production of each type of cement, local models were correctly selected to estimate specific consumption, resulting on a bank containing 3 energy consumption models. The estimation results showed very low estimation errors.

As a final observation to the work produced, it is possible to state that the context-based identification of energy consumption is a valid approach, which offers not only the ability to build more accurate energy consumption models, but also a structured way to identify how energy is used in industrial plants to track down causes of consumption.

An article was published describing the context-based approach proposed in this dissertation and was presented at the New Directions in Intelligent Decision Technologies

and Intelligent Interactive Multimedia Systems and Services Workshop during the KES-IDT-13 and KES-IIMSS-13 conference, that took place at Sesimbra, Portugal between 26-28 June 2013 [72] and attached in appendix A.3.

7.2 Future Work

There are several future work paths on different fields that can be followed from the contribution of this dissertation, such as context-based consumption decision support systems (intelligent decision support systems field), context-based control systems (control field) and context-based inter-process learning (adaptive systems field).

Over the context-identification and context-based consumption estimation developed in this dissertation, additional modules extending this work architecture, would be able to perform further consumption analysis based on the context identification and model's simulation. For instance, by putting together companies production goals and strategies with context-based simulation analysis, would allow, for instance, the forecasting of maximum consumption for a certain operating context, hence suitable for identify future peaks of consumption (peak load consumption problems), enabling companies to re-define and evaluate its strategies to avoid this type of situation. Furthermore, other consumption indicators may be derived, such as predict total, average and minimum consumption for each operating context. A suggestion of context-based energy analysis is depicted in figure 7.1 and a suggested achievement is illustrated on figure 7.2.

Incorporating context-based analysis information with decision support systems, would lead to the development of context-based decision support systems, based on context consumption analysis to support decision making and optimization measures.

Another interesting work path is to build control systems based on the developed CBEM. Consider for instance that a context variable (e.g. temperature) can be controlled (e.g. HVAC systems). The CBEM would forecast the energy consumption for certain

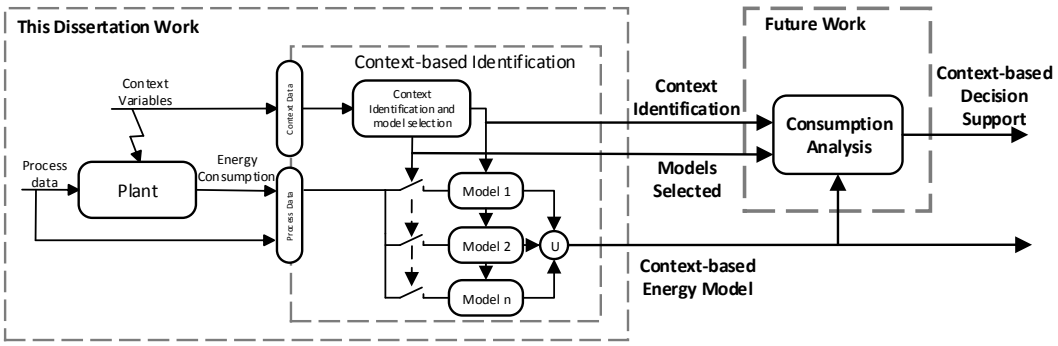

Figure 7.1: Future work path: Context-based consumption analysis systems.

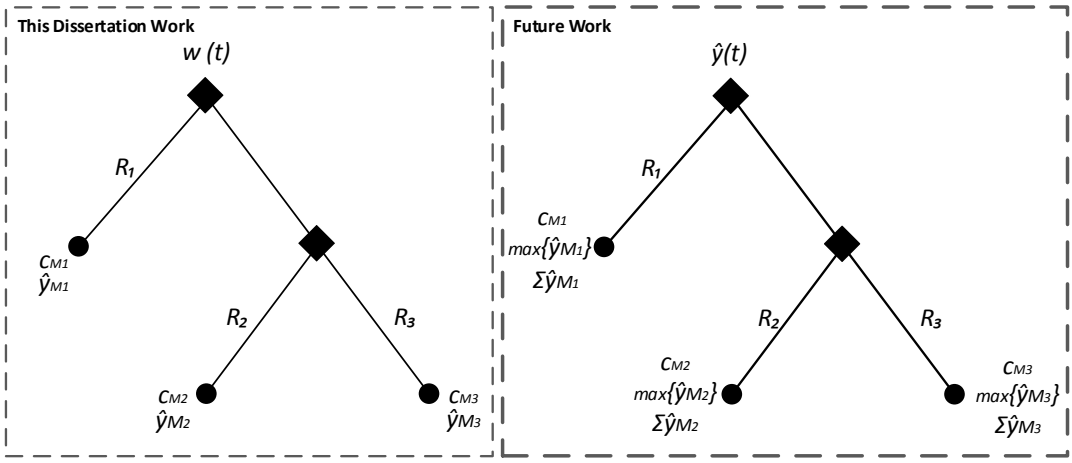

Figure 7.2: Future work path: Context-based consumption analysis results.

temperatures, allowing a predictive control over temperature before energy is actually consumed, hence entering in the field of control systems.

Also, CBEM may be used to reduce consumption between processes that perform a similar task. Consider two separate production chains with the same type of processes. By learning how context influences the consumption of each individual process, comparisons are possible to be derived in order to identify the best context to reduction consumption (adaptive systems).

The future application of context-based identification shows to be very promising in several fields and the possibilities are almost unlimited, mostly due to the evolution of technology and how it is directly related to the evolution of these fields. Hence, I hope this dissertation will become the starting point of future context-based systems and that its contribution enriches the knowledge and information of the research community and science.

References

[1] U. E. I. Administration, "International energy outlook 2014," Available: www.eia.gov/forecasts/ieo/industrial.cfm, September 2014, acessed: September 2014.

[2] C. V. Le, C. K. Pang, and O. P. Gan, "Energy saving and monitoring technologies in manufacturing systems with industrial case studies," *Industrial Electronics and Applications (ICIEA), 2012 7th IEEE Conference*, pp. 1916–1921, July 2012.

[3] D. Gielen and M. Taylor, "Modelling industrial energy use: The {IEAs} energy technology perspectives," *Energy Economics*, vol. 29, no. 4, pp. 889 – 912, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140988307000126

[4] R. Sahasrabudhe, P. Sistu, G. Sardar, and R. Gopinath, "Control and optimization in cement plants," *Control Systems, IEEE*, vol. 26, no. 6, pp. 56–63, Dec 2006.

[5] Y. Hongliang, J. Shaohong, and W. Xiaohong, "Research on control of cement clinker sintering process oriented to energy saving," *Control Conference (CCC), 2010 29th Chinese*, pp. 5726–5730, July 2010.

[6] A. Farzanegan, E. Ghasemi Ardi, A. Valian, and V. Hasanzadeh, "Simulation of clinker grinding circuits of cement plant based on process models calibrated using ga search method," *Journal of Central South University*, vol. 21, no. 2, pp. 799–810, 2014. [Online]. Available: http://dx.doi.org/10.1007/s11771-014-2003-7

[7] A. Beck and N. Jazdi, "Model-based electrical energy analysis of industrial automation systems," *2010 IEEE International Conference on Automation Quality and Testing Robotics (AQTR)*, vol. 2, pp. 1–6, May 2010.

[8] A. Beck and P. Gohner, "Analysis levels for improved user-centric energy cost analysis," *Information Systems and Technologies (CISTI), 2012 7th Iberian Conference*, pp. 1–6, June 2012.

[9] T. P. J. J.Heilala, K.Klobut, "Energy use parameters for energy efficiency enhancement in discrete manufacturing process," *7th CIRP International Conference on Intelligent Computation in Manufacturing Engineering*, June 2010.

[10] J. Heilala, K. Klobut, T. Salonen, P. Siltanen, R. Ruusu, A. Armijo, M. Sorli, L. Urosevic, P. Reimer, T. Fatur, Z. Gantar, and A. Jung, "Ambient intelligence based monitoring and energy efficiency optimisation system," *Assembly and Manufacturing (ISAM), 2011 IEEE International Symposium*, pp. 1–6, May 2011.

[11] S. Bandini and F. Sartori, "Improving the effectiveness of monitoring and control systems exploiting knowledge-based approaches," *Personal Ubiquitous Comput.*, vol. 9, no. 5, pp. 301–311, Sep. 2005. [Online]. Available: http://dx.doi.org/10.1007/s00779-004-0334-3

[12] E. Abdelaziz, R. Saidur, and S. Mekhilef, "A review on energy saving strategies in industrial sector," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 1, pp. 150 – 168, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S136403211000290X

[13] S. Mohagheghi and N. Raji, "Dynamic demand response solution for industrial customers," *Industry Applications Society Annual Meeting, 2013 IEEE*, pp. 1–9, Oct 2013.

[14] S. Rafiei and A. Bakhshai, "A review on energy efficiency optimization in smart grid," *IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society*, pp. 5916–5919, Oct 2012.

[15] J. Contreras, A. Losi, M. Russo, and F. Wu, "Simulation and evaluation of optimization problem solutions in distributed energy management systems," *Power Systems, IEEE Transactions*, vol. 17, no. 1, pp. 57–62, Feb 2002.

[16] M. Azadeh, S. Ghaderi, A. Guitiforooz, and M. Saberi, "Improved estimation of electricity demand function by integration of fuzzy system and data mining approach," *Industrial Technology, 2006. ICIT 2006. IEEE International Conference*, pp. 2160–2165, Dec 2006.

[17] F. Genova, F. Bellifemine, M. Gaspardone, M. Beoni, A. Cuda, and G. P. Fici, "Thermal and energy management system based on low cost wireless sensor network technology, to monitor, control and optimize energy consumption," *Telecommunication - Energy Special Conference (TELESCON), 2009 4th International Conference*, pp. 1–8, May 2009.

[18] L.-M. Wu and B.-S. Chen, "Modeling of energy efficiency indicator for semiconductor industry," *Industrial Engineering and Engineering Management, 2007 IEEE International Conference*, pp. 822–826, Dec 2007.

[19] H. Liu, Q. Zhao, W. Cao, N. Huang, and X. Zhao, "Simulation based evaluation and optimization for energy consumption of a typical welding shop," *Automation Science and Engineering (CASE), 2011 IEEE Conference*, pp. 660–665, Aug 2011.

[20] N. Mandic, H. Glavas, and I. Petrovic, "Influence of temperature drop on power and gas systems - analysis," *Energy Conference (ENERGYCON), 2014 IEEE International*, pp. 621–626, May 2014.

[21] C.-L. Hor, S. Watson, and S. Majithia, "Analyzing the impact of weather variables on monthly electricity demand," *Power Systems, IEEE Transactions*, vol. 20, no. 4, pp. 2078–2085, Nov 2005.

[22] D. Fortsch, "Wear impacts in slag grinding in various grinding technologies," *Cement Industry Technical Conference, Conference Record*, pp. 177–191, May 2005.

[23] V. Vakiloroaya, B. Samali, J. Madadnia, and Q. Ha, "Component-wise optimization for a commercial central cooling plant," *IECON 2011 - 37th Annual Conference on IEEE Industrial Electronics Society*, pp. 2769–2774, Nov 2011.

[24] M. Wallace, R. McBride, S. Aumi, P. Mhaskar, J. House, and T. Salsbury, "Energy efficient model predictive building temperature control," *Chemical Engineering Science*, vol. 69, no. 1, pp. 45 – 58, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0009250911004854

[25] X. Wang and V. Syrmos, "Nonlinear system identification and fault detection using hierarchical clustering analysis and local linear models," *Control Automation, 2007. MED '07. Mediterranean Conference*, pp. 1–6, June 2007.

[26] H. xiang Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3586 – 3592, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032112001438

[27] L. G. Swan and V. I. Ugursal, "Modeling of end-use energy consumption in the residential sector: A review of modeling techniques," *Renewable and Sustainable Energy Review*, vol. 13, no. 8, pp. 1819 – 1835, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032108001949

[28] G. K. Tso and K. K. Yau, "Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks," *Energy*, vol. 32, no. 9, pp. 1761 – 1768, 2007.

[29] S. Aman, Y. Simmhan, and V. K. Prasanna, "Improving energy use forecast for campus micro-grids using indirect indicators," *International Workshop on Domain Driven Data Mining (DDDM)*, pp. 1–9, 2011. [Online]. Available: http://ceng.usc.edu/~simmhan/pubs/aman-dddm-2011.pdf

[30] B. Dong, C. Cao, and S. E. Lee, "Applying support vector machines to predict building energy consumption in tropical region," *Energy and Buildings*, vol. 37, no. 5, pp. 545 – 553, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378778804002981

[31] V. Kecman and J. Brooks, "Locally linear support vector machines and other local models," *Neural Networks (IJCNN), The 2010 International Joint Conference*, pp. 1–6, July 2010.

[32] R. Winter, A. Boulanger, R. Anderson, and L. Wu, "Using Support Vector Machine to Forecast Energy Usage of a Manhattan Skyscraper," *AGU Fall Meeting Abstracts*, p. C984, Dec. 2011.

[33] X. C. Xi, A. N. Poo, and S. K. Chou, "Support vector regression model predictive control on a {HVAC} plant," *Control Engineering Practice*, vol. 15, no. 8, pp. 897 – 908, 2007, special Section on Modelling and Control for Participatory Planning and Managing Water Systems {IFAC} workshop on Modelling and Control for Participatory Planning and Managing Water Systems. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0967066106001766

[34] Z. Li, "Support vector machine model based predictive pid control system for cement rotary kiln," *Control and Decision Conference (CCDC), 2010 Chinese*, pp. 3117–3121, May 2010.

[35] Z. Jiu-Sun, G. Chuan-Hou, and P. Wei, "Modeling of high dimensional blast furnace system by manifold learning," *Control Conference (CCC), 2010 29th Chinese*, pp. 3157–3161, July 2010.

[36] W. Yi and L. Ying, "Applying ls-svm to predict primary energy consumption," *E-Product E-Service and E-Entertainment (ICEEE), 2010 International Conference*, pp. 1–4, Nov 2010.

[37] A. H. Neto and F. A. S. Fiorelli, "Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption," *Energy and Buildings*, vol. 40, no. 12, pp. 2169 – 2176, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378778808001448

[38] S. Wang and X. Dong, "Predicting china's energy consumption using artificial neural networks and genetic algorithms," *Business Intelligence and Financial Engineering, 2009. BIFE '09. International Conference*, pp. 8–11, July 2009.

[39] T. Yamada and T. Yabuta, "Remarks on neural network controller using different sigmoid functions," *Neural Networks, 1994. IEEE World Congress on Computational Intelligence., 1994 IEEE International Conference*, vol. 4, pp. 2628–2632 vol.4, Jun 1994.

[40] "Modeling of the appliance, lighting, and space-cooling energy consumptions in the residential sector using neural networks," *Applied Energy*, vol. 71, no. 2, pp. 87 – 110, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261901000496

[41] G. Escrivá-Escrivá, C. Álvarez Bel, C. Roldán-Blay, and M. Alcázar-Ortega, "New artificial neural network prediction method for electrical consumption forecasting based on building end-uses," *Energy and Buildings*, vol. 43, no. 11, pp. 3112 – 3119, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378778811003422

[42] C. Roldán-Blay, G. Escrivá-Escrivá, C. Álvarez Bel, C. Roldán-Porta, and J. Rodríguez-García, "Upgrade of an artificial neural network prediction method for electrical consumption forecasting using an hourly temperature curve model," *Energy and Buildings*, vol. 60, no. 0, pp. 38 – 46, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378778812006585

[43] Y. Bo, L. Yi, and Q. Shouning, "A rule-based cement kiln control system using neural networks," *Intelligent Processing Systems, 1997. ICIPS '97. 1997 IEEE International Conference*, vol. 1, pp. 493–497 vol.1, Oct 1997.

[44] Z. Xue and Z. Li, "Control system design of cement rotary kiln based on pid neuron networks," *Test and Measurement, 2009. ICTM '09. International Conference*, vol. 1, pp. 21–24, Dec 2009.

[45] M. Pusnik, B. Sucic, A. Podgornik, F. Al-Mansour, and T. Vuk, "Net fitting based production planning and decision support system for energy intensive industries," *Energy Conference (ENERGYCON), 2014 IEEE International*, pp. 1236–1242, May 2014.

[46] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*, 2nd ed. Cambridge, MA, USA: MIT Press, 1992.

[47] X. Yang, S. Guo, and H. Yang, "The establishment of energy consumption optimization model based on genetic algorithm," *Automation and Logistics, 2008. ICAL 2008. IEEE International Conference*, pp. 1426–1431, Sept 2008.

[48] A. Azadeh, R. Moghadam, S. Ghaderi, S. Tarverdian, and M. Saberi, "Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption in energy intensive sector," *E-Learning in Industrial Electronics, 2006 1ST IEEE International Conference*, pp. 58–63, Dec 2006.

[49] H. Zhao, R. Liu, Z. Zhao, and C. Fan, "Analysis of energy consumption prediction model based on genetic algorithm and wavelet neural network," *Intelligent Systems and Applications (ISA), 2011 3rd International Workshop*, pp. 1–4, May 2011.

[50] Z. Zhen-Ya, C. Hong-Mei, and Z. Shu-Guang, "An approach to the identification of temperature in intelligent building based on feed forward neural network and genetic algorithm," *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference*, vol. 2, pp. 286–290, July 2010.

[51] Z. Yuan, Z. Liu, and R. Pei, "Fuzzy control of cement raw meal production," *Automation and Logistics, 2008. ICAL 2008. IEEE International Conference*, pp. 1619–1624, Sept 2008.

[52] S.-M. Chen and L.-W. Lee, "Fuzzy decision-making based on likelihood-based comparison relations," *Fuzzy Systems, IEEE Transactions*, vol. 18, no. 3, pp. 613–628, June 2010.

[53] Z. Li, "Design of fuzzy neural network based control system for cement rotary kiln," *Informatics in Control, Automation and Robotics (CAR), 2010 2nd International Asia Conference*, vol. 2, pp. 290–293, March 2010.

[54] W. S. Rea, M. Reale, C. Cappelli, and J. Brown, "Identification of changes in mean with regression trees: An application to market research," *Econometric Reviews*, vol. 29, no. 5-6, pp. 754–777, 2010. [Online]. Available: http://EconPapers.repec.org/RePEc:taf:emetrv:v:29:y:2010:i:5-6:p:754-777

[55] M. Grigoletto, F. Lisi, and S. Petrone, *Complex Models and Computational Methods in Statistics*, ser. Contributions to Statistics. Milan, Italy: Springer, 2013, vol. 7.

[56] C. Huang and J. R. G. Townshend, "A stepwise regression tree for nonlinear approximation: applications to estimating subpixel land cover," *International Journal of Remote Sensing*, vol. 24, pp. 75–90, 2003.

[57] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shar, "Geometry from a Time Series," *Phys. Rev. Lett. 45*, p. 712, 1980.

[58] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*, ser. Classics in Applied Mathematics. New Jersey, USA: Society for Industrial and Applied Mathematics, 1974, vol. 15.

[59] J. Jiang and Y. Zhang, "A revisit to block and recursive least squares for parameter estimation," *Computers and Electrical Engineering*, vol. 30, no. 5, pp. 403 – 416, 2004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0045790604000217

[60] C. Lupu, D. Popescu, B. Ciubotaru, C. Petrescu, and G. Florea, "Switching solution for multiple-models control systems," *Control and Automation, 2006. MED '06. 14th Mediterranean Conference*, pp. 1–6, June 2006.

[61] M. Sanquer, F. Chatelain, M. El-Guedri, and N. Martin, "A smooth transition model for multiple-regime time series," *Signal Processing, IEEE Transactions*, vol. 61, no. 7, pp. 1835–1847, April 2013.

[62] D. Shangodoyin, O. Adebile, and R. Arnab, "Smooth transition modelling: empirical investigation of logistic and exponential transition functions." *Model Assist. Stat. Appl.*, vol. 3, no. 1, pp. 43–57, 2008.

[63] G. W. K. Colman and J. Wells, "On the use of rls with covariance reset in tracking scenarios with discontinuities," *Electrical and Computer Engineering, 2006. CCECE '06. Canadian Conference*, pp. 693–696, May 2006.

[64] A. Jankovic, W. Valery, and E. Davis, "Cement grinding optimisation," *Minerals Engineering*, vol. 17, no. 11–12, pp. 1075 – 1081, 2004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0892687504002043

[65] J. Satterfield, "Advancements in meeting new regulations with activated carbon: Results from full scale cement plant field tests," *Cement Industry Technical Conference (CIC), 2013 IEEE-IAS/PCA*, pp. 1–10, April 2013.

[66] D. Tsamatsoulis, "Modeling of raw materials blending in raw meal grinding systems," *Proceedings of the 14th WSEAS International Conference on Systems: Part of the 14th WSEAS CSCC Multiconference*, vol. 1, pp. 129–134, 2010. [Online]. Available: http://dl.acm.org/citation.cfm?id=1984140.1984170

[67] L. B. Palma, J. C. Cruz, P. S. Gil, and F. V. Coito, "Simulator of dynamical systems and pid control based on java language," *International Journal of Online Engineering (iJOE)*, vol. 9, no. S8, pp. 59–61, December 2013. [Online]. Available: http://online-journals.org/index.php/i-joe/article/view/3386

[68] L. Brito Palma, J. Costa Cruz, F. Vieira Coito, and P. Sousa Gil, "Java-based simulator of dynamical systems and pid control," *Experiment@ International Conference (exp.at'13), 2013 2nd*, pp. 72–77, Sept 2013.

[69] L. Brito Palma, J. Costa Cruz, P. Sousa Gil, and F. Vieira Coito, "Java stand-alone application for linear control systems simulation," *Remote Engineering and Virtual Instrumentation (REV), 2014 11th International Conference*, pp. 361–369, Feb 2014.

[70] L. Brito Palma, J. Costa Cruz, F. Vieira Coito, and P. Sousa Gil, "Interactive demonstration of a java-based simulator of dynamical systems," *Experiment@ International Conference (exp.at'13), 2013 2nd*, pp. 176–177, Sept 2013.

[71] L. Brito Palma, J. Costa Cruz, P. Sousa Gil, and F. Vieira Coito, "Dynamical systems simulator based on a java application," *Remote Engineering and Virtual Instrumentation (REV), 2014 11th International Conference*, pp. 397–398, Feb 2014.

[72] J. Cruz, R. Neves-Silva, and M. Marques, "Context-based identification of Energy Consumption in Industrial Plants," *Advances in Smart Systems Research*, vol. 3, no. 1, pp. 31–37, 2014. [Online]. Available: http://nimbusvault.net/publications/koala/assr/208.html

Appendix

A.1 Cement Mill Data of November and December 2012

Figure A.1 shows the raw data during the month of November.

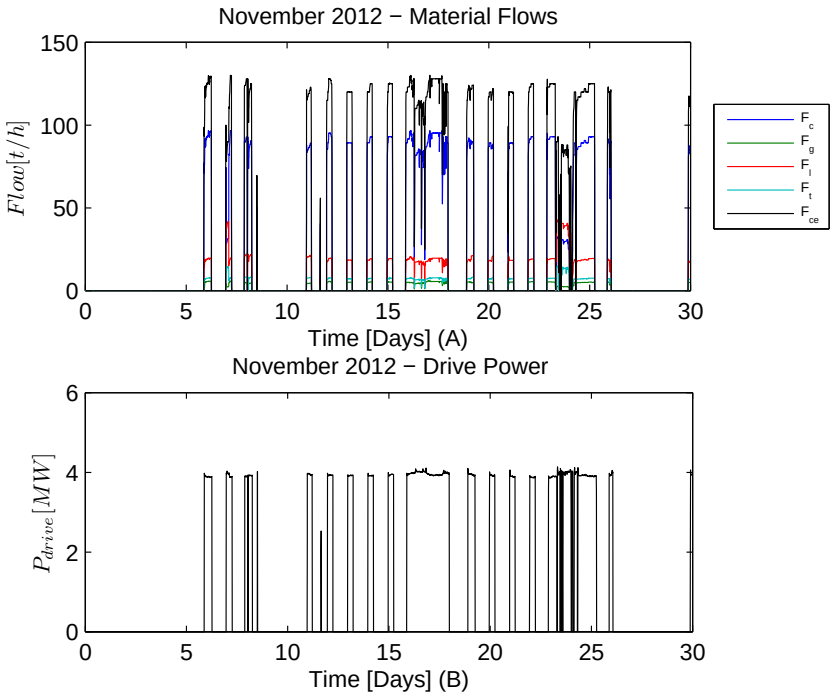

Figure A.1: November 2012 raw data. (A) Flow of input and output materials. (B) Power consumption.

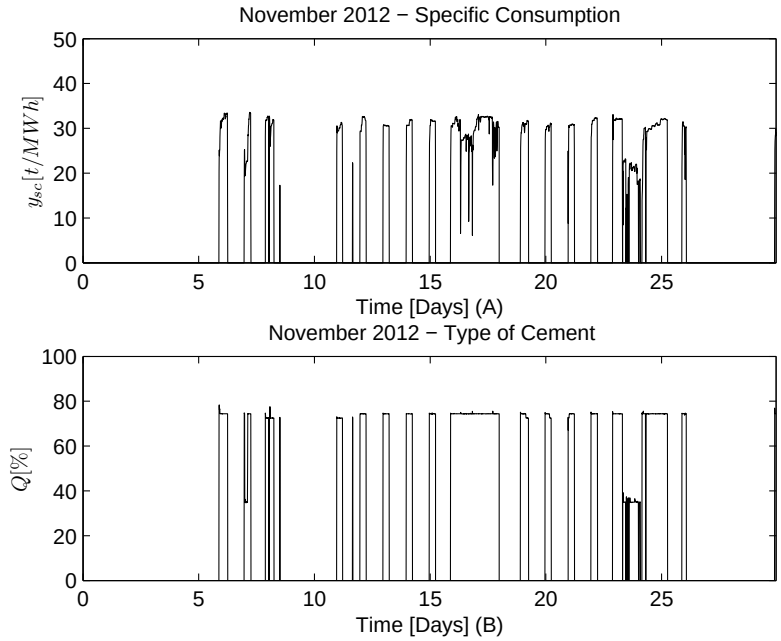

Figure A.2: November 2012 - Cement specific consumption and cement type. (A) Specific consumption. (B) Type of cement.

From figure A.1 (A) it is shown the individual flow of materials entering the mill and the production flow of cement. In (B) the power consumption of the cement mill drive is shown. Particularly around the 24th day of production, the ratio between clinker and other materials that compose the cement to be milled is largely different than of the rest of the month. In this period, the cement flow rate is also much lower. This pattern was already observed on the previous months for cement type B production. Although not so evident, cement type B is also produced on the 7th day for only 40 minutes. On the rest of the month, cement type A was produced.

From figure A.2 (B) is observed that cement type B produced during the periods mentioned on the previous paragraph, as the average ratio of clinker is nearly 35%. The cement type A is identified by a ratio of approximately 75%. Again, a difference in the specific consumption is observed when cement type B is under production as depicted in figure A.2 (A).

Regarding the month of December 2012, it is shown the raw data on figure A.3. During this month, the production of type B cement is again noticed, this time occurring between the 7th and the 8th day, whereas type A cement is produced on the remaining month.

Specific consumption and type of cement are depicted on figure A.4, respectively in (A) and (B) for the month of December.

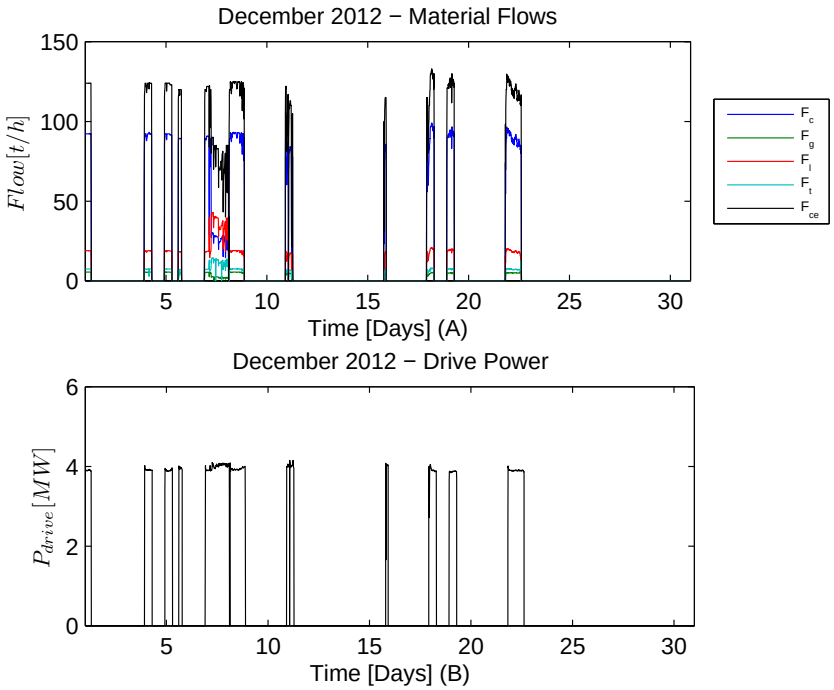

Figure A.3: December 2012 raw data. (A) Flow of input and output materials. (B) Power consumption.

As it is expected, from figure A.4 analysis, a large drop on the cement rate production is observed when cement type B is being produced. This is the same energy consumption footprint that has been observed in the previous months.

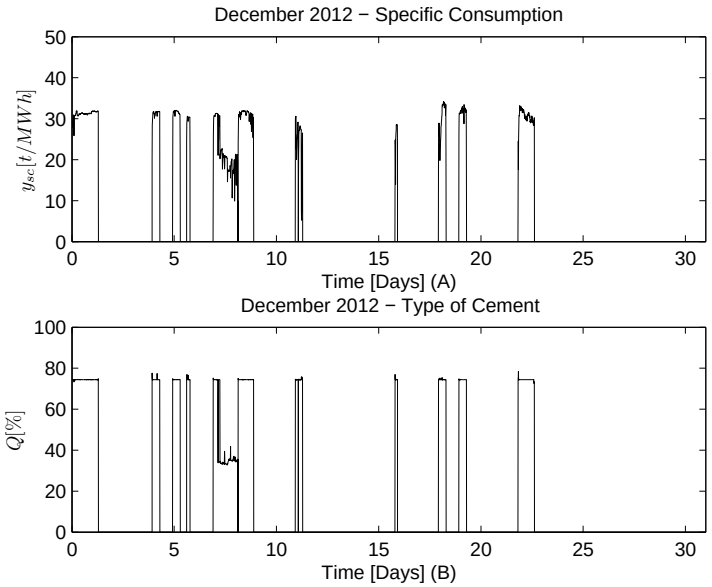

Figure A.4: December 2012 - Cement specific consumption and cement type. (A) Specific consumption. (B) Type of cement.

Table A.1 depicts the main information retrieved from the cement operation during November and December.

Table A.1: Data Summary for the months of November and December.

Data	November 2012	December 2012
Type of Cement	A ; B	A ; B
Clinker Ratio [%]	75 ; 35	75 ; 35
Production Rate [t/h]	119.2 ; 89.3	118.1 ; 87.6
Total Production [kt]	22.3	16.7
Specific Consumption [t/MWh]	30.1 ; 23.3	30.2 ; 22.3
Shortest Operating Time [min]	40	120

A.2 CBEM Experimental Results of November and December 2012

Figure A.5 shows the CBEM system performance on identifying context changes for the full month of November. As it is depicted in (A) the regression tree algorithm is able to successfully distinguish between different types of cementa, allowing the bank of models to select which model should be active for estimation in shown in (B).

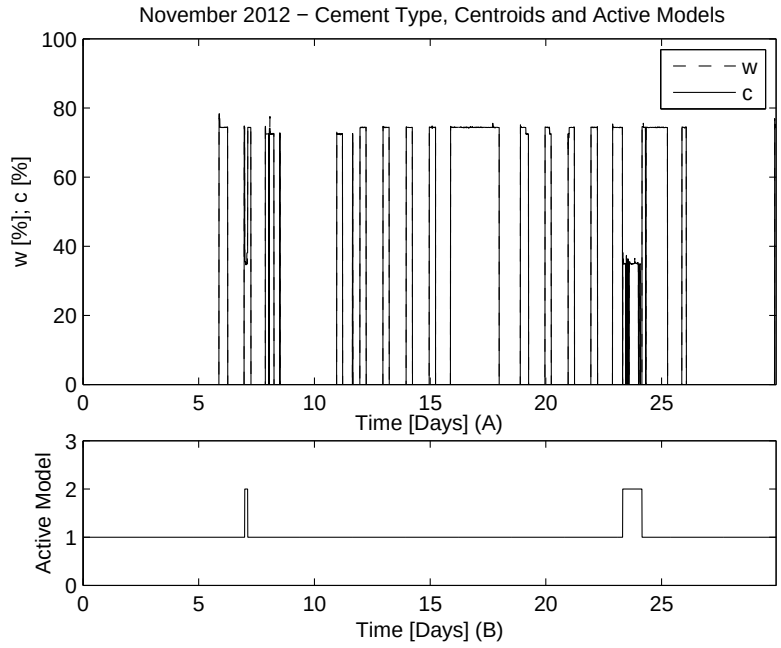

Figure A.5: CBEM context identification results of November 2012 - (A) Type of cement and centroids. (B) Active model.

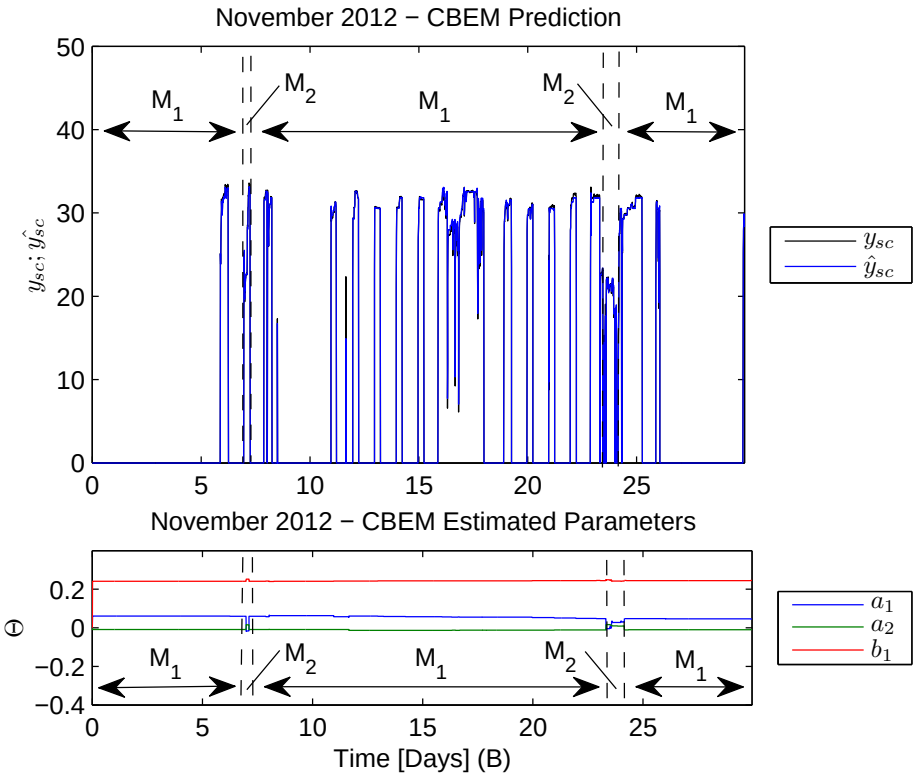

Figure A.6: CBEM prediction results of November 2012 - (A) Specific consumption prediction (B) Estimated Parameters.

The CBEM system prediction of specific consumption is depicted in figure A.6 (A), highlighting the regions where the model estimated for cement type A and B where selected to be active for prediction, allowing an accurate prediction with average absolute estimation error lower than 0.09%. In (B) is depicted how parameters are estimated over time.

Tables A.2 and A.3 show respectively, the resultant bank of models for this month and the numerical result for the average absolute estimation error.

Table A.2: Models Estimated in November 2012.

Cement Type	**Model**	**Centroid [%]**	**Estimated Parameters**
A	M_1	$c_{M_1} = 74.61$	$\Theta_{M_1} = [0.0452\ ;\ 0.0093\ ;\ 0.2454]$
B	M_2	$c_{M_2} = 35.59$	$\Theta_{M_2} = [0.0304\ ;\ 0.0096\ ;\ 0.2418]$

Table A.3: CBEM absolute average estimation error for November 2012.

CBEM Absolute Average Estimation Error $\overline{\lvert\Delta\epsilon\rvert}$
$\overline{\lvert\Delta\epsilon\rvert} = 0.088\, t/MWh$

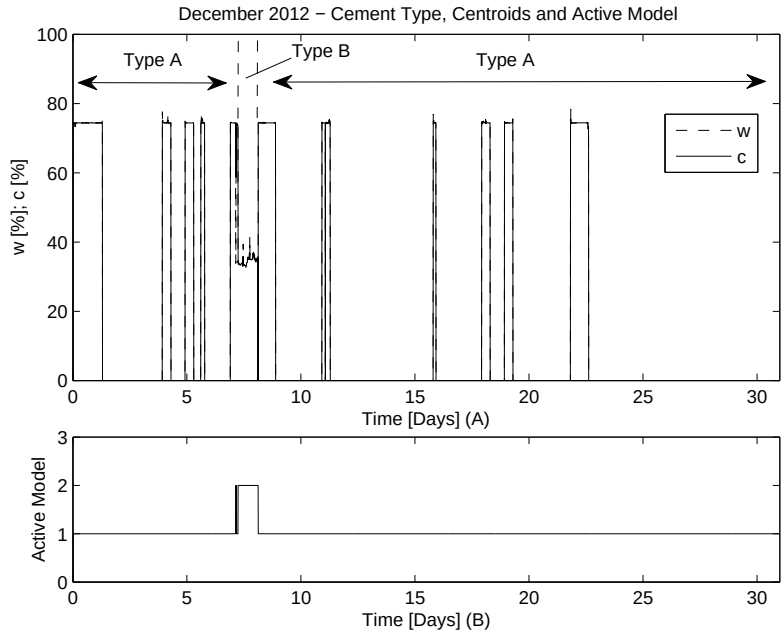

Figure A.7: CBEM context identification results of December 2012 - (A) Type of cement and centroids. (B) Active model.

Regarding the month of December 2012, type A and type B cement are produced. Figure A.7 (A) depicts the CBEM system results on evaluating when the operating context changes, from type A to type B cement. In (B) it is shown the model selected for estimation, respectively model M_1 for type A cement production and M_2 for type B.

Figure A.8 (A) depicts the CBEM prediction of specific consumption, in (B) the estimated parameters and in (C) a zoom on the region where type B cement is under production. Again, the average absolute estimation error is less than 0.09 t/MWh, which confirms an accurate prediction and context identification.

Tables A.4 and A.5 show the bank of models state on the end of this month, and numerical results of the average absolute estimation error.

Table A.4: Models Estimated in December 2012.

Cement Type	**Model**	**Centroid [%]**	**Estimated Parameters**
A	M_1	$c_{M_1} = 74.61$	$\Theta_{M_1} = [0.0485\ ;\ 0.0065\ ;\ 0.2540]$
B	M_2	$c_{M_2} = 35.59$	$\Theta_{M_2} = [0.0312\ ;\ 0.0059\ ;\ 0.2416]$

Table A.5: CBEM absolute average estimation error for December 2012.

CBEM Absolute Average Estimation Error $\overline{\lvert\Delta\epsilon\rvert}$
$\overline{\lvert\Delta\epsilon\rvert} = 0.0752\ t/MWh$

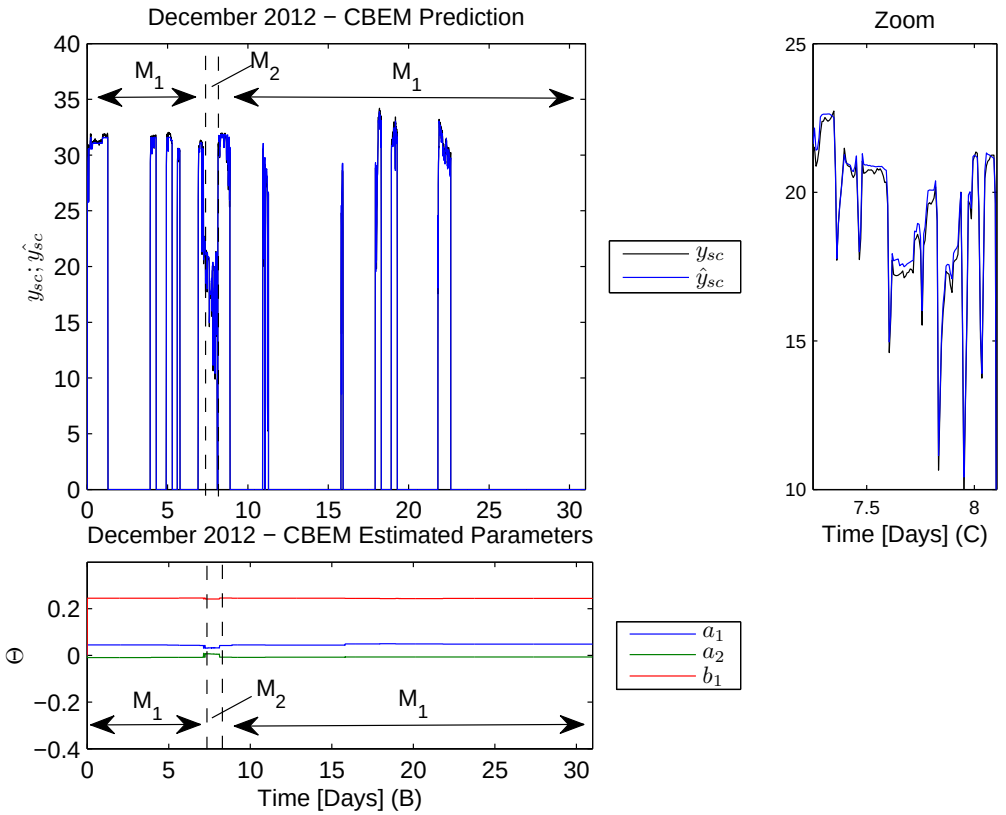

Figure A.8: CBEM prediction results of December 2012 - (A) Specific consumption prediction (B) Estimated parameters. (C) Zoom on type B cement operating regime.

A.3 Published Paper

Advances in Smart Systems Research : ISSN 2050-8662 : http://nimbusvault.net/publications/koala/assr/
Vol. 3. No. 2 : pp.31-37 : iswk13-007

Context-based Identification of Energy Consumption in Industrial Plants

João Cruz [a], Rui Neves-Silva [a], Maria Marques[b]
[a] *DEE-FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal*
jm.cruz@campus.fct.unl.pt
[b] *UNINOVA, FCT Campus, 2829-516 Caparica, Portugal*

Abstract. In industrial environment, plants are exposed to multiple contextual factors, which affect the way that energy is used. Actual energy models provide guidelines for energy consumption, but fail in explaining some patterns observed in practice. This paper provides a novel approach for building energy models of industrial plants, based on the influence of context variables. Multiple RLS algorithms are used for multi-model context-based estimation of energy consumption. For context identification, regression trees are used to divide the context variables into regions, where the local models are estimated. The effectiveness of the proposed approach is illustrated with some examples and validated with a practical case in a real industrial cement plant.

Keywords. Energy Consumption; Context Variables; RLS Multi-Model Estimation; Regression Trees;

Introduction

In an industrial environment, energy consumption depends on multiple factors, such as production factors (e.g. type of product being manufactured; equipment condition), ambient factors (e.g. humidity of the air; year season) or human factors (e.g. operators experience). Actual energy consumption models are based on fixed operating regimes, disregarding context awareness. These models can provide guidelines for energy use, but tend to fail in explaining some behaviors. Also, recently technological advances in electronics and communication networks, allows the continuous and distributed monitoring of contextual information in industrial environments, which makes it easier to acquire information. The main contribution of this work is a novel approach for building energy consumption models based on context variables.
While an industrial plant is performing a specific task, it is exposed simultaneously to multiple contextual variables. As the task is being performed, variables related to the plant operation, also affect the energy consumption. From this point of view, these two types of variables are at the same level. Yet, context variables are classified as external variables that are kept steady for periods of time that go beyond of plant time constants, whereas plant operation variables are defined as being the input, directly related to the task that is being performed, e.g. while a milling is grinding, the flux of raw material is considered as a plant input variable, while the type of cement being produced is a context variable. The steadiness period, suggests the partition of the context variables into several regions in which a local model can be estimated. Therefore, the

construction of context-based energy models involves two entirely different stages. In a first stage, the context variables are partitioned into several regions, using regression trees, as they deal well with continuously measured data, and provide fast algorithms for computation [1]. And the second stage is the energy consumption modeling. Some modeling techniques based on regression analysis[2], decision trees[3] and neural networks[4] are suggested by different authors. Although, the modeling is made by one unique nonlinear descriptor. As different regions of the contexts variables affect differently the energy consumption, a multi-model approach is proposed, using multiple RLS algorithms for model estimation. Therefore, for each region a local energy consumption model is estimated.

1. Context-based Identification Concept

As mentioned earlier, the context-based identification of energy consumption involves two different stages. The global architecture is depicted in Fig. 1. The block named Region and Centroid Identification is responsible for the partitioning of context variables into a set of regions; and calculation of respective centroids. For model estimation, the same data regressor is used. The global nonlinear energy consumption model is obtained by joining all the local models estimated for each region of the context variables.

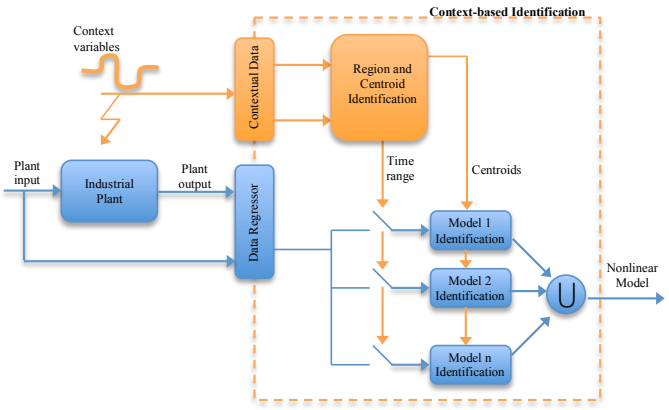

Figure 1 – Context based identification architecture.

2. Context Identification

This chapter describes how context variables are divided into several regions, in which local models are used to estimate the energy consumption. The proposed approach starts by splitting the context variables into several regions, where a centroid is calculated. For this purpose, regression trees are used for each context variable. Fig. 2 depicts the global architecture.
Each of the tree leaves defines a region, that is a time range for model estimation, and has attached to it a real value which applies only in that region, named centroid.
Existing methods for tree construction utilize different splitting and pruning approaches, such as distance measures, likelihood-ratio tests or least squares criterion [5].
The regression trees applied in this paper, use the least squares criterion, and are constructed by recursively splitting leaf nodes in order to maximize the tree homogeneity. Also, a steadiness time must be granted to able the models to be identified; this is achieved by stopping the tree from growing when the region is less than a specified dwell time.
The centroid identifies where an estimated model is valid. I.e. suppose that the samples between k_i and k_f are all samples belonging to the same region R of a context variable w. Then the centroid c_w^R is the mean of w over R, as described by the following equation

$$c_w^R = \frac{1}{k_f - k_i} \sum_{i=k_i}^{k_f} w(i) \tag{1}$$

Therefore, a model is only valid in the proximity of the centroid, delimited by the region boundaries.
After all context variables are partitioned, it is necessary to intersect the regions and join the respective centroids. The Time Ranges and Centroids Matcher block performs this function.

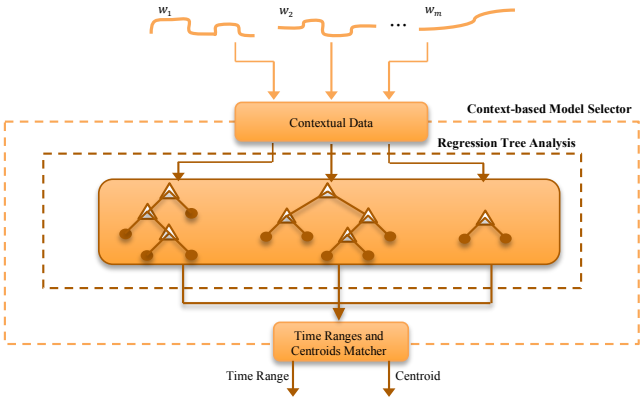

Figure 2. General architecture for context-based model selection.

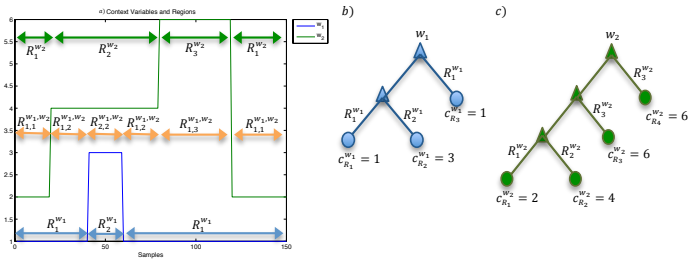

Figure 3. a) Context variables and regions; The green, blue and orange arrows mark the regions of w_1, w_2 and its intersection, respectively; b) Regression tree of w_1; c) Regression tree of w_2.

For example, consider two distinct context variables w_1 and w_2, that represent the worker's ID number operating simultaneously the same machine, as depicted in fig. X a). The respective regression trees are represented in fig. X b) and c); and the numerical values of centroids and time-ranges for each region are presented in Table 1 and 2.

Table 1. Region, Centroid and Time Range results of regression tree analysis for context variable w_1.

Region	Centroid	Time Range
$R_1^{w_1}$	$c_{R_1}^{w_1} = 1$	$k < 40.5 \cup k \geq 59.5$
$R_2^{w_1}$	$c_{R_2}^{w_1} = 3$	$k \geq 40.5 \cap k < 59.5$

Table 2. Region, Centroid and Time Range results of regression tree analysis for context variable w_2.

Region	Centroid	Time Range
$R_1^{w_2}$	$c_{R_1}^{w_2} = 2$	$k < 19.5$
$R_2^{w_2}$	$c_{R_2}^{w_2} = 4$	$k \geq 19.5 \cap k < 79.5$
$R_3^{w_2}$	$c_{R_3}^{w_2} = 6$	$k \geq 79.5$

As the time-ranges of the regression trees overlap, it's necessary to intersect the regions and join the centroids, so that the active model in that region can be uniquely identified. Therefore, a new set of regions is created and presented in table 3.
By this approach, the estimated energy model for each region captures the influence of multiple contexts, which is coherent to what happens in an industrial environment where the plant is exposed at the same time to all the contextual variables, and all influence simultaneously the plant's energy consumption.

Table 3. Results of intersected regions and joined centroids.

Intersected Regions	Joined Centroids
$R_{1,1}^{w_1w_2} = R_1^{w_1} \cap R_1^{w_2}$	$c_{R_{1,1}}^{w_1w_2} = (1,2)$
$R_{1,2}^{w_1w_2} = R_1^{w_1} \cap R_2^{w_2}$	$c_{R_{1,2}}^{w_1w_2} = (1,4)$
$R_{2,2}^{w_1w_2} = R_2^{w_1} \cap R_2^{w_2}$	$c_{R_{2,2}}^{w_1w_2} = (3,6)$
$R_{1,3}^{w_1w_2} = R_1^{w_1} \cap R_3^{w_2}$	$c_{R_{1,3}}^{w_1w_2} = (1,6)$

Context-based identification of Energy Consumption in Industrial Plants
João Cruz, Rui Neves-Silva, Maria Marques

3. Energy Models Estimation

After defining the centroids and time range for each region of the context variables, local modeling is used to estimate the plants energy consumption. There are as many models as regions; and while a model is being estimated, all the others are freezed. Therefore, while a plant is performing a specific task, e.g. milling, a RLS algorithm is used per region, in order to capture the energy consumption model. Also, the combination of multiple local models, defines the plant's global energy consumption.
In this outline, consider that the plant's energy consumption, can be modeled by the following equation

$$y(k) = \theta^T \varphi(k) + e(k) \tag{2}$$

Where $y(k)$ is the energy consumption; θ is a vector of unknown parameters; $\varphi(k)$ is a regression vector of continuously measurements of past energy consumption and plant input, according to the task that is being performed; $e(k)$ is white noise disturbance; and k is the time discrete independent variable. Fig. 4 a) depicts the measured energy consumption signal; and Fig. 4 b) the measured process input u, e.g. flux of raw material.
Also, consider that the context variables w_1 and w_2, exemplified in the previous chapter and again represented in Fig. 4 c), influence directly the parameters θ. Thus, as θ is context-based dependent, eq. 2 must be re-written as follows

$$y(k) = \theta^T(w) \cdot \varphi(k) + e(k) \tag{3}$$

Where w is a vector of context variables measurements.
The main goal of the RLS algorithms is to estimate the unknown context-based parameters as the data is continuously measured. Consider that for each time discrete sample, the regression vector $\varphi(k)$ is described by

$$\varphi(k) = [y(k-1)\,;\, y(k-2)\,;\, y(k-3)\,;\, u(k-1)\,;\, u(k-2)] \tag{4}$$

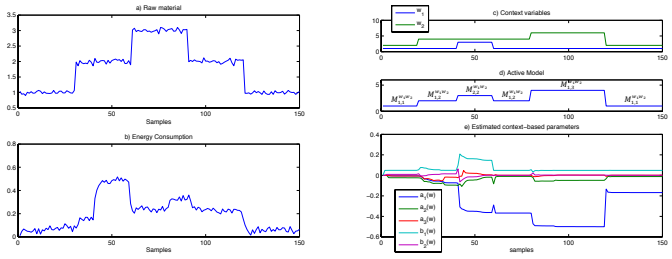

Figure 4. a) Raw material as input signal; b) Energy consumption as output signal; c) Context variables; d) Estimated context-based parameters; e) Active model

Therefore, the unknown context-based parameters being estimated by each RLS algorithm is defined by

$$\hat{\theta} = [a_1(w)\ ;\ a_2(w)\ ;\ a_3(w)\ ;\ b_1(w)\ ;\ b_2(w)]^T \tag{5}$$

Fig. 4 d) depicts the active model being estimated, and Fig. 4 e) the RLS parameter estimation for each model, whose numerical values are presented in table 4.
Notice that as the RLS algorithm takes more measurements into account, the parameters converge to a final value. This states that the model is as more accurate as more measurements are made. Also, as the region of the context variable changes, the parameters seems to be strongly affected, which suggest that both dynamical and static properties of the energy model are being influenced by the context variables. As the static property is defined by the steady state relation, for the given example the relation between flux of raw material and energy consumption, it can be stated that the context is affecting the way that the energy is being used, in terms of production rate.

Table 4. Model, centroid and respective context-based parameter estimation.

Model	Centroid	Estimated Parameters
$M_{1,1}^{W_1W_2}$	$c_{R_{1,1}}^{W_1W_2} = (1,2)$	[-0.1667 ; -0.0127 ; 0.0013 ; 0.0489 ; 0.0006]
$M_{1,2}^{W_1W_2}$	$c_{R_{1,2}}^{W_1W_2} = (1,4)$	[-0.3670 ; -0.0117 ; 0.0026 ; 0.0488 ; 0.0003]
$M_{2,2}^{W_1W_2}$	$c_{R_{2,2}}^{W_1W_2} = (3,6)$	[-0.3612 ; -0.0166 ; 0.0040 ; 0.1432 ; 0.0047]
$M_{1,3}^{W_1W_2}$	$c_{R_{1,3}}^{W_1W_2} = (1,6)$	[-0.4978 ; -0.0511 ; 0.0032 ; 0.0487 ; 0.0029]

For model evaluation, the estimated models are used for simulation of the energy consumption, with the same identical setup. Figure 5 depicts the comparison between the measured energy consumption y, and the simulation $\tilde{y}_m$ using the multi-model context-based approach. Fig. 5 also depicts a simulation $\tilde{y}_s$, based on the classical method of using a unique model also estimated by RLS algorithm.
The mean square error of the multi-model context-based approach is 5.2802×10^{-4} while the mean square error of the classic method is 1.1172×10^{-2}, which indicates a very promising and applicable concept for improving the actual used energy consumption models.

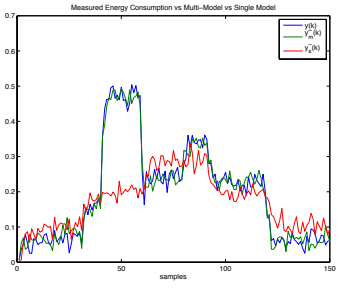

Figure 5. Comparison between measured energy consumption, multi-model and single model simulation.

4. Conclusion

Context-based energy consumption models can be used for better understanding of the influence of context variables in industrial plants, but also for extrapolating the energy consumption into the future, according to different scenarios. Intelligent decision support system, based on these models, will be able to improve production strategies, reduce energy costs, and solve other type of problems such as peak load management. The context-based identification concept has been validated by the results obtained in the cement plant, that shown very accurate model's prediction.